

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 09:22:18 ; Search time 165 Seconds
(without alignments)
480.520 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047
Sequence: 1 MSKKKGLSAEKRTMMEIF.....FEENKIDRTFGIPEDFYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1047	100.0	205	7	ADDB4537 121PI1 P
2	1047	100.0	205	7	ADJ70015 Human hea
3	1047	100.0	205	8	ADM83851 Human can
4	1047	100.0	205	8	ADM83810 Human can
5	1047	100.0	205	8	ADM83812 Human can
6	1047	100.0	205	8	ADM83793 Human can
7	1047	100.0	205	8	ADM83811 Human can
8	1047	100.0	205	8	ADM83804 Human can
9	1047	100.0	219	6	ABP75541 Human sec
10	1036.5	99.0	206	8	ADM83834 Human can
11	1036.5	99.0	206	8	ADM83835 Human can
12	1011	96.6	198	8	ADM83814 Human can
13	975	93.1	190	4	AAW40043 Human can
14	975	93.1	190	7	ADDB4547 121PI1 v
15	975	93.1	190	8	ADM83857 Human can
16	975	93.1	190	8	ADM83859 Human can
17	975	93.1	190	8	ADM83803 Human can
18	975	93.1	190	8	ADM83809 Human can
19	975	93.1	190	8	ADM83858 Human can
20	948	90.5	205	8	ADM83813 Human can
21	945.5	90.3	190	7	ADDB4545 121PI1 v
22	945.5	90.3	190	8	ADM83808 Human can
23	945.5	90.3	190	8	ADM83801 Human can
24	945.5	90.3	190	8	ADM83853 Human can
25	945.5	90.3	190	8	ADM83852 Human can

26	614	58.6	122	8	ADM83843 Human can
27	606	57.9	122	4	AAW41829 Human pol
28	591	56.4	122	7	ADDB4543 121PI1 v
29	591	56.4	122	8	ADM83799 Human can
30	591	56.4	122	8	ADM83845 Human can
31	591	56.4	122	8	ADM83807 Human can
32	591	56.4	122	8	ADM83844 Human can
33	584	55.8	119	7	ADDB4541 121PI1 v
34	584	55.8	119	8	ADM83797 Human can
35	584	55.8	119	8	ADM83806 Human can
36	584	55.8	119	8	ADM83837 Human can
37	465	44.4	126	7	ADDB4539 121PI1 v
38	465	44.4	126	8	ADM83836 Human can
39	465	44.4	126	8	ADM83795 Human can
40	465	44.4	126	8	ADM83805 Human can
41	349	33.3	200	8	ADM83815 Human can
42	329.5	31.5	182	4	ABG12241 Novel hyp
43	287	27.4	79	4	ABG12341 Novel hum
44	255.5	24.4	218	5	ADH32846 Yeast smo
45	135.5	12.9	1281	8	ADP26646 Mouse dyn

ALIGNMENTS

RESULT 1	ADDB4537	standard; protein; 205 AA.
ID	ADDB4537	
AC	ADDB4537;	
XX		
DT	29-JAN-2004	(first entry)
XX		
DE	121PI1 protein.	
XX		
KW	121PI1; 121PI1 modulation; human; chromosome 4q; cytostatic;	
XX	gene therapy; vaccine; cancer; immune response; immunisation.	
OS	Homo sapiens.	
XX		
PN	WO200295009-A2.	
XX		
PD	28-NOV-2002.	
XX		
PF	28-FEB-2002; 2002MO-US006242.	
XX		
PR	05-MAR-2001; 2001US-00799250.	
XX		
PA	(AGEN-) AGENSYS INC.	
XX		
PI	Challita-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;	
XX	Jakovovits A;	
PI	WPI; 2003-156757/15.	
XX	N-PSDB; ADDB4536.	
DR		
PT	Composition comprising a substance that modulates the status of 121PI1,	
XX	useful in diagnosing, preventing, prognosticating or treating patients	
PT	with cancer that expresses 121PI1, such as breast, colon, ovarian or	
XX	lung cancer.	
PS	Claim 19; Fig 2A; 285pp; English.	
XX		
CC	The present invention describes a composition (I) comprising a substance	
XX	that modulates the status of 121PI1 (gene and encoded protein), or a	
CC	molecule that is modulated by 121PI1, where the status of a cell that	
XX	expresses 121PI1 is modulated. The human 121PI1 gene maps to chromosome	
CC	4q. (I) has cytostatic activity, and can be used in gene therapy, and in	
XX	vaccines. The composition (I) can be used for diagnosing, preventing,	
CC	prognosticating or treating patients with cancer that expresses 121PI1,	
XX	such as breast, colon, ovarian or lung cancer. The 121PI1 gene or its	
CC	fragment can be used to elicit a humoral or cellular immune response.	
XX	121PI1 antibodies can be used in active or passive immunisation. 121PI1	

CC polynucleotides are useful as probes and primers for the amplification or
CC detection of 121P1F1 genes, as coding sequences for directing the
CC expression of 121P1F1 polypeptides, or as tools for modulating or
CC inhibiting the expression of 121P1F1 genes. The present sequence is used
CC in the exemplification of the present invention.

XX Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 7; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.5e-88;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKRTMMEIFSETKDVFOQLDKLEKTAPEKGTAMSVKVLQSLVDDGVY 60
DB 1 MSKKKGLSAEKKRTMMEIFSETKDVFOQLDKLEKTAPEKGTAMSVKVLQSLVDDGVY 60
QY 61 DCEKIGTSNYWAFPSKALHARKHKLVLVLSQSLSGSKHSLQKSIKAKIGRCETEE 120
DB 61 DCEKIGTSNYWAFPSKALHARKHKLVLVLSQSLSGSKHSLQKSIKAKIGRCETEE 120
QY 121 TRIAKELSSLRDQREQLKAEVEKXKDCDPQVVEIRQANKVAKKANRWTDNIFAISWA 180
DB 121 TRIAKELSSLRDQREQLKAEVEKXKDCDPQVVEIRQANKVAKKANRWTDNIFAISWA 180
QY 181 KRRKFGPEENKIDRTFGIPEDFDYID 205
DB 181 KRRKFGPEENKIDRTFGIPEDFDYID 205

RESULT 2
ADJ70015

ID ADJ70015 standard; protein; 205 AA.

AC ADJ70015;

XX 06-MAY-2004 (first entry)

DE Human heart mitochondrial protein as a therapeutic target SeqID1821.

XX mitochondrial; human; screening assay; diabetes mellitus;
KM Huntington's disease; osteoarthritis;
KM Leber's hereditary optic neuropathy; LHON;
KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KM myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KM neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
KM osteopathic; ophthalmological; cytoskeletal.

XX Homo sapiens.

XX WO2003087768-A2.

XX 23-OCT-2003.

XX 04-APR-2003; 2003WO-US010870.

XX 12-APR-2002; 2002US-0372843P.

XX 17-JUN-2002; 2002US-038987P.

XX 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

XX (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
XX Warnock DE;

XX WPI; 2003-845369/78.

XX Identifying a mitochondrial target for drug screening assays and for
XX treating diseases associated with altered mitochondrial function,
XX comprises detecting a modified polypeptide in a sample and correlating
XX with the disease.

XX Claim 1; SEQ ID NO 1821; 180pp; English.

XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic, and
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytoskeletal activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 7; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.5e-88;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKRTMMEIFSETKDVFOQLDKLEKTAPEKGTAMSVKVLQSLVDDGVY 60
DB 1 MSKKKGLSAEKKRTMMEIFSETKDVFOQLDKLEKTAPEKGTAMSVKVLQSLVDDGVY 60
QY 61 DCEKIGTSNYWAFPSKALHARKHKLVLVLSQSLSGSKHSLQKSIKAKIGRCETEE 120
DB 61 DCEKIGTSNYWAFPSKALHARKHKLVLVLSQSLSGSKHSLQKSIKAKIGRCETEE 120
QY 121 TRIAKELSSLRDQREQLKAEVEKXKDCDPQVVEIRQANKVAKKANRWTDNIFAISWA 180
DB 121 TRIAKELSSLRDQREQLKAEVEKXKDCDPQVVEIRQANKVAKKANRWTDNIFAISWA 180
QY 181 KRRKFGPEENKIDRTFGIPEDFDYID 205
DB 181 KRRKFGPEENKIDRTFGIPEDFDYID 205

RESULT 3
ADM83851

ID ADM83851 standard; protein; 205 AA.

AC ADM83851;

XX 03-JUN-2004 (first entry)

DE Human cancer gene 121P1F1 protein #6.

XX Human; cancer gene 121P1F1; cytoskeletal; cancer; chromosome 4q; HLA;
KM human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
KM colon cancer; lung cancer; pancreatic cancer; breast cancer;
KM cervical cancer; stomach cancer; gene therapy; vaccine.

XX Homo sapiens.

XX US2003223997-A1.

XX 04-DEC-2003.

XX 28-FEB-2002; 2002US-00087190.

XX 08-FEB-2001; 2001US-00779250.

XX (CHAL/) CHALLITA-RID P M.

XX (HUBER/) HUBERT R S.

XX (RAIT/) RAITANO A B.

XX (FARI/) FARIS M.

XX (AFAR/) AFAR D B H.

XX (GEWM/) GE W.

XX (JAKO/) JAKOBOVITS A.

PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
PI Jakobovits A;
XX
XX WPI; 2004-060522/06.
XX
XX
XX New composition comprising a substance that modulates the status of
PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
PT lung cancer.

Example 5; SEQ ID NO 61; 211pp; English.

The invention relates to a composition comprising a substance that
modulates the status of 121P1F1 (a protein encoded by a cancer expressed
gene) or a molecule that is modulated by 121P1F1 where status of the cell
that expresses 121P1F1 is modulated. Also included are a pharmaceutical
composition comprising the novel composition in a human unit dose form, a
recombinant protein comprising an antigen-binding region of a monoclonal
antibody, a non-human transgenic animal that produces an antibody, a
hybridoma that produces an antibody, a single chain monoclonal antibody
that immunospecifically binds to a 121P1F1-related protein (comprising
the variable domains of the heavy and light chains of a monoclonal
antibody), a vector comprising a polynucleotide that encodes a single
chain monoclonal antibody, a polynucleotide that encodes an analogue
peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
that expresses 121P1F1, inhibiting growth of cancer cells that expresses
121P1F1 (comprising administering to the cells the composition), treating
a patient who bears cancer cells that express 121P1F1, generating a
mammalian immune response directed to 121P1F1, inducing an immune
response, monitoring 121P1F1 gene products in a biological sample from a
patient who has or who is suspected of having cancer, monitoring the
presence of cancer in an individual and an assay for detecting the
presence of a 121P1F1-related protein or polynucleotide in a biological
sample from a patient who has or who is suspected of having cancer. The
composition may comprise a polynucleotide that comprises a 121P1F1-
related protein coding sequence provided that the coding sequence does
not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
substance also comprises a polynucleotide that encodes at least one
peptide given in 16 Tables (given in the specification), the peptides
being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its
splice variants. The composition is useful for detecting, treating or
preventing cancer, preferably prostate cancer, bladder cancer, kidney
cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
cervical cancer or stomach cancer. The composition can also be used as a
vaccine to treat or prevent cancer that expresses or overexpresses
CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
CC sequence is a 121P1F1 protein (full-length or fragment).

Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 8; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.5e-88;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRRKGGSLAEKRTMWEIIFSETKOVFOIKLEKAPKKGTTAMSVKVLSDVDDGV 60
DB 1 MRRKGGSLAEKRTMWEIIFSETKOVFOIKLEKAPKKGTTAMSVKVLSDVDDGV 60
QY 61 DERIGTSTNYVAFPSKALHARKHKLVEESQSESGQKASIOKSIERAKIGRCETTER 120
DB 61 DERIGTSTNYVAFPSKALHARKHKLVEESQSESGQKASIOKSIERAKIGRCETTER 120
QY 121 TRLAKELSLRDQREOLKAEVEKYKCDPQVVEEIRQANKVAKAANRWTDNIPAIKSWA 180
DB 121 TRLAKELSLRDQREOLKAEVEKYKCDPQVVEEIRQANKVAKAANRWTDNIPAIKSWA 180
QY 181 KKKFGFEENKIDRTFCGIPEDFDYID 205
DB 181 KKKFGFEENKIDRTFCGIPEDFDYID 205

RESULT 4
ADM83810

ID ADM83810 standard; protein; 205 AA.
XX
XX
AC ADM83810;
XX
XX 03-JUN-2004 (first entry)
DT
XX
XX Human cancer gene 121P1F1 protein #3.
DE
XX
XX Human; cancer gene 121P1F1, cytostatic; cancer; chromosome 4q; HLA;
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
KW cervical cancer; stomach cancer; gene therapy; vaccine.
XX
XX Homo sapiens.
OS
XX
XX US2003223997-A1.
XX
XX
XX 04-DEC-2003.
XX
XX 28-FEB-2002; 2002US-00087190.
XX
XX 08-FEB-2001; 2001US-00779250.
XX
XX (CHAL//) CHALLITA-EID P M.
PA (HUBE//) HUBERT R S.
PA (RAIT//) RAITANO A B.
PA (FAIR//) FARIS M.
PA (AFAR//) AFAR D E H.
PA (GEW//) GE W.
PA (JAKO//) JAKOBOVITS A.
XX
PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
PI Jakobovits A;
XX
XX WPI; 2004-060522/06.
XX
XX
XX New composition comprising a substance that modulates the status of
PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
PT lung cancer.
XX
XX
XX Disclosure; SEQ ID NO 20; 211pp; English.
XX
XX
XX The invention relates to a composition comprising a substance that
modulates the status of 121P1F1 (a protein encoded by a cancer expressed
gene) or a molecule that is modulated by 121P1F1 where status of the cell
that expresses 121P1F1 is modulated. Also included are a pharmaceutical
composition comprising the novel composition in a human unit dose form, a
recombinant protein comprising an antigen-binding region of a monoclonal
antibody, a non-human transgenic animal that produces an antibody, a
hybridoma that produces an antibody, a single chain monoclonal antibody
that immunospecifically binds to a 121P1F1-related protein (comprising
the variable domains of the heavy and light chains of a monoclonal
antibody), a vector comprising a polynucleotide that encodes a single
chain monoclonal antibody, a polynucleotide that encodes an analogue
peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
that expresses 121P1F1, inhibiting growth of cancer cells that expresses
121P1F1 (comprising administering to the cells the composition), treating
a patient who bears cancer cells that express 121P1F1, generating a
mammalian immune response directed to 121P1F1, inducing an immune
response, monitoring 121P1F1 gene products in a biological sample from a
patient who has or who is suspected of having cancer, monitoring the
presence of cancer in an individual and an assay for detecting the
presence of a 121P1F1-related protein or polynucleotide in a biological
sample from a patient who has or who is suspected of having cancer. The
composition may comprise a polynucleotide that comprises a 121P1F1-
related protein coding sequence provided that the coding sequence does
not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
substance also comprises a polynucleotide that encodes at least one
peptide given in 16 Tables (given in the specification), the peptides
being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its
splice variants. The composition is useful for detecting, treating or
preventing cancer, preferably prostate cancer, bladder cancer, kidney

CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
CC cervical cancer or stomach cancer. The composition can also be used as a
CC vaccine to treat or prevent cancer that expresses or overexpresses
CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
CC sequence is a 121P1F1 protein (full-length or fragment).
XX
SQ Sequence 205 AA;
Query Match 100.0%; Score 1047; DB 8; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.5e-88;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKKGLSABEKKTRMMEIFSETKQVFOKDLKIAPEKGTITANSVKVTLQSLVDDGV 60
DB 1 MSKKKGLSABEKKTRMMEIFSETKQVFOKDLKIAPEKGTITANSVKVTLQSLVDDGV 60
QY 61 DCEIRIGTSNYMAFPESKALHARKHKLVELESQSESGQKASLOKSTIEKAKIGRCETER 120
DB 61 DCEIRIGTSNYMAFPESKALHARKHKLVELESQSESGQKASLOKSTIEKAKIGRCETER 120
QY 121 TRIAKELSSLRDQREQLKAVEREYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSMA 180
DB 121 TRIAKELSSLRDQREQLKAVEREYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSMA 180
QY 181 KRKGFEBENKIDRTFGIPEDFDYID 205
DB 181 KRKGFEBENKIDRTFGIPEDFDYID 205
RESULT 5
ADM83812 standard; protein; 205 AA.
AC ADM83812;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human cancer gene 121P1F1 protein #5.
XX
KW Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
KW cervical cancer; stomach cancer; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN US2003223997-A1.
XX
PD 04-DEC-2003.
XX
PF 28-FEB-2002; 2002US-00087190.
XX
PR 08-FEB-2001; 2001US-00779250.
XX
PA (CHAL/) CHALLITA-ETD P M.
PA (HUBE/) HUBERT R S.
PA (RAIT/) RAITANO A B.
PA (FARI/) FARIS M.
PA (AFAR/) AFAR D E H.
PA (GEWM/) GE W.
PA (JAKO/) JAKOBOVITS A.
XX
XX Chailita-Ed PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,
PI Jakobovits A;
XX
XX WPI; 2004-060522/06.
XX
XX
XX New composition comprising a substance that modulates the status of
PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
PT lung cancer.
XX
PS Disclosure; SEQ ID NO 22; 211pp; English.

XX
XX The invention relates to a composition comprising a substance that
CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed
CC gene) or a molecule that is modulated by 121P1F1 where status of the cell
CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical
CC composition comprising the novel composition in a human unit dose form, a
CC recombinant protein comprising an antigen-binding region of a monoclonal
CC antibody, a non-human transgenic animal that produces an antibody, a
CC hybridoma that produces an antibody, a single chain monoclonal antibody
CC that immunospecifically binds to a 121P1F1-related protein (comprising
CC the variable domains of the heavy and light chains of a monoclonal
CC antibody), a vector comprising a polynucleotide that encodes a single
CC chain monoclonal antibody, a polynucleotide that encodes an analogue
CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses
CC 121P1F1 (comprising administering to the cells the composition), treating
CC a patient who bears cancer cells that express 121P1F1, generating a
CC mammalian immune response directed to 121P1F1, inducing an immune
CC response, monitoring 121P1F1 gene products in a biological sample from a
CC patient who has or who is suspected of having cancer, monitoring the
CC presence of cancer in an individual and an assay for detecting the
CC presence of a 121P1F1-related protein or polynucleotide in a biological
CC sample from a patient who has or who is suspected of having cancer. The
CC composition may comprise a polynucleotide that comprises a 121P1F1-
CC related protein coding sequence provided that the coding sequence does
CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
CC substance also comprises a polynucleotide that encodes at least one
CC peptide given in 16 Tables (given in the specification), the peptides
CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its
CC splice variants. The composition is useful for detecting, treating or
CC preventing cancer, preferably prostate cancer, bladder cancer, kidney
CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
CC cervical cancer or stomach cancer. The composition can also be used as a
CC vaccine to treat or prevent cancer that expresses or overexpresses
CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
CC sequence is a 121P1F1 protein (full-length or fragment).
XX
SQ Sequence 205 AA;
Query Match 100.0%; Score 1047; DB 8; Length 205;
Best Local Similarity 100.0%; Pred. No. 1.5e-88;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKKGLSABEKKTRMMEIFSETKQVFOKDLKIAPEKGTITANSVKVTLQSLVDDGV 60
DB 1 MSKKKGLSABEKKTRMMEIFSETKQVFOKDLKIAPEKGTITANSVKVTLQSLVDDGV 60
QY 61 DCEIRIGTSNYMAFPESKALHARKHKLVELESQSESGQKASLOKSTIEKAKIGRCETER 120
DB 61 DCEIRIGTSNYMAFPESKALHARKHKLVELESQSESGQKASLOKSTIEKAKIGRCETER 120
QY 121 TRIAKELSSLRDQREQLKAVEREYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSMA 180
DB 121 TRIAKELSSLRDQREQLKAVEREYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSMA 180
QY 181 KRKGFEBENKIDRTFGIPEDFDYID 205
DB 181 KRKGFEBENKIDRTFGIPEDFDYID 205
RESULT 6
ADM83793 standard; protein; 205 AA.
AC ADM83793;
XX
XX
XX
XX 03-JUN-2004 (first entry)
XX
XX
XX Human cancer gene 121P1F1 protein #1.
XX
XX Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
KW colon cancer; lung cancer; pancreatic cancer; breast cancer;

CC the variable domains of the heavy and light chains of a monoclonal
 CC antibody), a vector comprising a polynucleotide that encodes a single
 CC chain monoclonal antibody, a polynucleotide that encodes an analogue
 CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
 CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses
 CC 121P1F1 (comprising administering to the cells the composition), treating
 CC a patient who bears cancer cells that express 121P1F1, generating a
 CC mammalian immune response directed to 121P1F1, inducing an immune
 CC response, monitoring 121P1F1 gene products in a biological sample from a
 CC patient who has or who is suspected of having cancer, monitoring the
 CC presence of cancer in an individual and an assay for detecting the
 CC presence of a 121P1F1-related protein or polynucleotide in a biological
 CC sample from a patient who has or who is suspected of having cancer. The
 CC composition may comprise a polynucleotide that comprises a 121P1F1-
 CC related protein coding sequence provided that the coding sequence does
 CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
 CC substance also comprises a polynucleotide that encodes at least one
 CC peptide given in 16 Tables (given in the specification), the peptides
 CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its
 CC splice variants. The composition is useful for detecting, treating or
 CC preventing cancer, preferably prostate cancer, bladder cancer, kidney
 CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
 CC cervical cancer or stomach cancer. The composition can also be used as a
 CC vaccine to treat or prevent cancer that expresses or overexpresses
 CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
 CC sequence is a 121P1F1 protein (full-length or fragment).

XX SQ Sequence 205 AA;

CC Query Match 100.0%; Score 1047; DB 8; Length 205;
 CC Best Local Similarity 100.0%; Pred. No. 1.5e-88;
 CC Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKRTMMEIFSETKQVFOQLDKLEKIAPEKGTITAMSVKVLQSLVDDGW 60
 Db 1 MSKKKGLSAEKKRTMMEIFSETKQVFOQLDKLEKIAPEKGTITAMSVKVLQSLVDDGW 60

QY 61 DCRIGTSNYWAFPSKALHARKHKLVLLESQSLSEGSQKHAISQKSTIEKAKIGRCETERR 120
 Db 61 DCRIGTSNYWAFPSKALHARKHKLVLLESQSLSEGSQKHAISQKSTIEKAKIGRCETERR 120

QY 121 TRIAKELSLRDQREOLKAVERKYKDCDPQVVEEIRQANKVAKEANRWTDNIFAIKSWA 180
 Db 121 TRIAKELSLRDQREOLKAVERKYKDCDPQVVEEIRQANKVAKEANRWTDNIFAIKSWA 180

QY 181 KRRFGFENKIDRTFGIPEDFDYID 205
 Db 181 KRRFGFENKIDRTFGIPEDFDYID 205

RESULT 8
 ADM83804 ID ADM83804 standard; protein; 205 AA.
 XX ADM83804;
 AC
 XX 03-JUN-2004 (first entry)
 DT
 XX
 DB Human cancer gene 121P1F1 protein #2.
 XX
 KW Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
 KW cervical cancer; stomach cancer; gene therapy; vaccine.
 OS Homo sapiens.
 XX
 XX US2003223997-A1.
 PN
 XX 04-DEC-2003.
 PD
 XX 28-FEB-2002; 2002US-00087190.
 PF
 XX

PR 08-FEB-2001; 2001US-00779250.
 XX
 XX (CHAL/) CHALLITA-EID P. M.
 PA (HUBE/) HUBERT R. S.
 PA (RAIT/) RAITANO A. B.
 PA (FARI/) FARIS M.
 PA (AFAR/) AFAR D E H.
 PA (GEWM/) GE W.
 PA (JAKO/) JAKOBOVITS A.
 XX
 PI Chailita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
 XX Jakobovits A;
 XX WPI; 2004-060522/06.
 XX
 PT New composition comprising a substance that modulates the status of
 PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
 PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
 PT lung cancer.
 XX
 XX Disclosure; SEQ ID NO 14; 211P; English.
 XX
 XX The invention relates to a composition comprising a substance that
 CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed
 CC gene) or a molecule that is modulated by 121P1F1 where status of the cell
 CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical
 CC composition comprising the novel composition in a human unit dose form, a
 CC recombinant protein comprising an antigen-binding region of a monoclonal
 CC antibody, a non-human transgenic animal that produces an antibody, a
 CC hybridoma that produces an antibody, a single chain monoclonal antibody
 CC that immunospecifically binds to a 121P1F1-related protein (comprising
 CC the variable domains of the heavy and light chains of a monoclonal
 CC antibody), a vector comprising a polynucleotide that encodes a single
 CC chain monoclonal antibody, a polynucleotide that encodes an analogue
 CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
 CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses
 CC 121P1F1 (comprising administering to the cells the composition), treating
 CC a patient who bears cancer cells that express 121P1F1, generating a
 CC mammalian immune response directed to 121P1F1, inducing an immune
 CC response, monitoring 121P1F1 gene products in a biological sample from a
 CC patient who has or who is suspected of having cancer, monitoring the
 CC presence of cancer in an individual and an assay for detecting the
 CC presence of a 121P1F1-related protein or polynucleotide in a biological
 CC sample from a patient who has or who is suspected of having cancer. The
 CC composition may comprise a polynucleotide that comprises a 121P1F1-
 CC related protein coding sequence provided that the coding sequence does
 CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
 CC substance also comprises a polynucleotide that encodes at least one
 CC peptide given in 16 Tables (given in the specification), the peptides
 CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its
 CC splice variants. The composition is useful for detecting, treating or
 CC preventing cancer, preferably prostate cancer, bladder cancer, kidney
 CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
 CC cervical cancer or stomach cancer. The composition can also be used as a
 CC vaccine to treat or prevent cancer that expresses or overexpresses
 CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
 CC sequence is a 121P1F1 protein (full-length or fragment).

XX SQ Sequence 205 AA;

CC Query Match 100.0%; Score 1047; DB 8; Length 205;
 CC Best Local Similarity 100.0%; Pred. No. 1.5e-88;
 CC Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKRTMMEIFSETKQVFOQLDKLEKIAPEKGTITAMSVKVLQSLVDDGW 60
 Db 1 MSKKKGLSAEKKRTMMEIFSETKQVFOQLDKLEKIAPEKGTITAMSVKVLQSLVDDGW 60

QY 61 DCRIGTSNYWAFPSKALHARKHKLVLLESQSLSEGSQKHAISQKSTIEKAKIGRCETERR 120
 Db 61 DCRIGTSNYWAFPSKALHARKHKLVLLESQSLSEGSQKHAISQKSTIEKAKIGRCETERR 120

QY 121 TRIAKELSLRDQREOLKAVERKYKDCDPQVVEEIRQANKVAKEANRWTDNIFAIKSWA 180

Db 121 TRLAKELSLDQREQLKAEVEKXKDCDPQVEEIRQANKVAKKAAANWNTDNIFAIKSWA 180
 QY 181 KRKGFEENKIDRTFGIPEDFDYID 205
 Db 181 KRKGFEENKIDRTFGIPEDFDYID 205

RESULT 9

ABP75541
 ID ABP75541 standard; protein; 219 AA.

XX ABP75541;
 AC

XX 10-FEB-2003 (first entry)
 DT

XX Human secretory polypeptide SPTM SEQ ID NO 725.
 DE

XX Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
 KW anti-inflammatory; immunosuppressive; neuroprotective; nootropic;
 KW neuroleptic; anticonvulsant; cyostatic; antiparkinsonian; anxiolytic;
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
 KW secretory polynucleotide; secretory protein.

OS Homo sapiens.
 XX

XX WO200283876-A2.
 PN

XX 24-OCT-2002.
 PD

XX 27-MAR-2002; 2002MO-US009921.
 PF

XX 29-MAR-2001; 2001US-0280067P.
 PR

XX 16-MAY-2001; 2001US-0280068P.
 PR

XX 17-MAY-2001; 2001US-0291829P.
 PR

XX 17-MAY-2001; 2001US-0291849P.
 PR

XX 19-JUN-2001; 2001US-0299428P.
 PR

XX 20-JUN-2001; 2001US-0299776P.
 PR

XX 20-JUN-2001; 2001US-0300001P.
 PR

XX (INCY-) INCYTE GENOMICS INC.
 PA

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshy SR;
 PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstlin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Matwaha R, Lo A, Ian RY, Urashka ME;
 XX

XX WPI; 2003-075543/07.
 DR

XX N-PSDB; AB235987.
 DR

XX New human secretory proteins and polynucleotides, useful for diagnosing,
 PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
 PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
 PT cancers.

XX Claim 27; SEQ ID NO 725; 458bp + Sequence Listing; English.
 PS

XX The invention relates to a secretory polynucleotide (designated sptm)
 CC comprising any of 567 polynucleotide sequences (AB235837-AB236403), a
 CC naturally occurring polynucleotide sequence at least 90 % identical to
 CC the polynucleotide sequence, a polynucleotide complementary to them or an
 CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
 CC creating, preventing or diagnosing a disease or condition associated with
 CC the expression of functional SPTM. These are particularly useful for
 CC diagnosing, treating or preventing autoimmune/inflammatory disorders
 CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
 CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
 CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
 CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
 CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
 CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
 CC breast, cervix or prostate). The present sequence is one of the SPTM
 CC proteins of the invention (ABP75384-ABP75962). Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

XX Sequence 219 AA;
 SQ

Query Match 100.0%; Score 1047; DB 6; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1,7e-88;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGSLAEKRRIRMEIFSETQDVQKLEKIAPEKIGITMVSKEVQSLVDQGV 60
 Db 15 MSKKKGSLAEKRRIRMEIFSETQDVQKLEKIAPEKIGITMVSKEVQSLVDQGV 74
 QY 61 DCEKIGTNTYMAFPKSKALHARKKLEVLSEQLSEGSQKASLQKSIKAKIGCETEE 120
 Db 75 DCEKIGTNTYMAFPKSKALHARKKLEVLSEQLSEGSQKASLQKSIKAKIGCETEE 134
 QY 121 TRLAKELSLDQREQLKAEVEKXKDCDPQVEEIRQANKVAKKAAANWNTDNIFAIKSWA 180
 Db 135 TRLAKELSLDQREQLKAEVEKXKDCDPQVEEIRQANKVAKKAAANWNTDNIFAIKSWA 194
 QY 181 KRKGFEENKIDRTFGIPEDFDYID 205
 Db 195 KRKGFEENKIDRTFGIPEDFDYID 219

RESULT 10

ADM83834
 ID ADM83834 standard; protein; 206 AA.

XX ADM83834;
 AC

XX 03-JUN-2004 (first entry)
 DT

XX Human cancer gene 121P1F1 variant protein #1.
 DE

XX Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
 KW cervical cancer; stomach cancer; gene therapy; vaccine.

OS Homo sapiens.
 XX

XX US2003223997-A1.
 PN

XX 04-DEC-2003.
 XX

XX 28-FEB-2002; 2002US-00087190.
 PF

XX 08-FEB-2001; 2001US-00779250.
 PR

XX (CHAL/) CHALITA-BID P M.
 PA

XX (HUBE/) HUBERT R S.
 PA

XX (RAIT/) RAITANO A B.
 PA

XX (FARI/) FARIS M.
 PA

XX (AFAR/) AFAR D E H.
 PA

XX (GEWU/) GE W.
 PA

XX (JAKO/) JAKOBOVITS A.
 PI

XX Chalita-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
 PI Jakobovits A;
 XX

XX WPI; 2004-060522/06.
 DR

XX New composition comprising a substance that modulates the status of
 PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,

PT treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.

PS Example 5; SEQ ID NO 44; 211pp; English.

The invention relates to a composition comprising a substance that modulates the status of 121PIPI1 (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121PIPI1 where status of the cell that expresses 121PIPI1 is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIPI1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIPI1, inhibiting growth of cancer cells that expresses 121PIPI1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIPI1, generating a mammalian immune response directed to 121PIPI1, inducing an immune response, monitoring 121PIPI1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIPI1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIPI1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIPI1 (ADM83793). The substance also comprises a polynucleotide that encodes at least one peptide given in 1a Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIPI1 or its splicing variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIPI1. The gene for 121PIPI1 is located on chromosome 4q. The present sequence is a 121PIPI1 protein (full-length or fragment).

SQ Sequence 206 AA;

Query Match	99.0%	Score 1036.5	DB 8	Length 206
Best Local Similarity	99.5%	Pred. No. 1.4e-87		
Matches 205; Conservative	0	Mismatches 0	Indels 1	Gaps 1

Qy	1	MSKKKGLSAEKCTRMMEIFSEFKVDFOLKJDEKLAPEKKGITANSVEVYQSLVDDGMV	60
Db	1	MSKKKGLSAEKCTRMMEIFSEFKVDFOLKJDEKLAPEKKGITANSVEVYQSLVDDGMV	60
Qy	61	DCERIGTSNYNMAFPSKALHARKHKLVEVLS-QLSBGSQKHALSLQKSIEMAKIGRCETE	119
Db	61	DCERIGTSNYNMAFPSALHARKHKLVEVLSQQLSBSGSQKHALSLQKSIEMAKIGRCETE	120
Qy	120	RTRLAKGLSLRDQREOLKAVEKRYKDCDPVEEIRIQRANKYAKAANRWTDNIFAIKSW	179
Db	121	RTRLAKGLSLRDQREOLKAVEKRYKDCDPVEEIRIQRANKYAKAANRWTDNIFAIKSW	180
Qy	180	AKRKFGEENKIDRTFGIPEDPYID 205	
Db	181	AKRKFGEENKIDRTFGIPEDPYID 206	

RESULT 11

ID ADM83835 standard; protein; 206 AA.

AC ADM838357

DT 03-JUN-2004 (first entry)

DE Human cancer gene 121P1F1 variant protein #2.

XX Human: cancer gene 121P1P1; cytosolic; cancer; chromosome 4q; HLA;
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
 KW cervical cancer; stomach cancer; gene therapy; vaccine.

OS Homo sapiens.

PN US2003223997-A1.

PD 04-DEC-2003.

PF 28-FEB-2002; 2002US-00087190.

PR 08-FEB-2001; 2001US-00779250

PA (CHAL/) CHALLITA-EID P M

PA (RAIT/) RAITANO A B.

PA (AFAR/) AFAR D E H.

PA (JAKO/) JAKOBOVITS

PI Challita-Eid PM, H

XX

XX

PT 121PIF1 or a molecule that is modulated by 121PIF1, useful for detecting

PT lung cancer.

PS Example 5; SEQ ID NO 45; 211pp; English.

The invention relates to a composition comprising a substance that modulates the status of 121PIF1 (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121PIF1 where status of the cell that expresses 121PIF1 is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1, inhibiting growth of cancer cells that expresses 121PIF1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIF1, generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIF1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIF1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIF1 (A083793). The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIF1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present sequence is a 121PIF1 protein (full-length or fragment).

Sequence 206 AA:

Sequence 206 AA;

55

Query Match 99.0%; Score 1036.5; DB 8; Length 206;
 Best Local Similarity 99.5%; Pred. No. 1.4e-87;
 Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MSKKKGLSAEKKRTMMELFSTKQVFOFKDLKIAPKKGTITAMSVKVLQSLVDDGAV 60
 DB 1 MSKKKGLSAEKKRTMMELFSTKQVFOFKDLKIAPKKGTITAMSVKVLQSLVDDGAV 60
 QY 61 DEERIGTSNYWAFPSKALHARKHKLKLEVESQSLSGSQGHASLQKSIKAKIGRCETEE 119
 DB 61 DEERIGTSNYWAFPSKALHARKHKLKLEVESQSLSGSQGHASLQKSIKAKIGRCETEE 120
 QY 120 RTFLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKYAKKXANRWNTDNIPAIKSM 179
 DB 121 RTFLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKYAKKXANRWNTDNIPAIKSM 180
 QY 180 AKRKGFEENKIDRTFGIPEDFDYID 205
 DB 181 AKRKGFEENKIDRTFGIPEDFDYID 206

RESULT 12
 ADM83814 standard; protein; 198 AA.
 XX
 AC ADM83814;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Human cancer gene 121P1F1 protein fragment.
 XX
 KM Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
 KM human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
 KM colon cancer; lung cancer; pancreatic cancer; breast cancer;
 KM cervical cancer; stomach cancer; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN US2003223997-A1.
 XX
 PD 04-DEC-2003.
 XX
 PF 28-FEB-2002; 2002US-00087190.
 PR 08-FEB-2001; 2001US-00779250.
 XX
 PA (CHAL/) CHALLITA-EID P M.
 PA (HUBE/) HUBERT R S.
 PA (RAIT/) RAITANO A B.
 PA (FARI/) PARIS M.
 PA (AFAR/) AFAR D E H.
 PA (GEWU/) GE W.
 PA (JAKO/) JAKOBYITS A.
 XX
 PI Chalita-Eid PM, Hubert RS, Raitano AB, Paris M, Afar DEH, Ge W;
 PI Jakobovits A;
 XX
 DR WPI; 2004-060522/06.
 XX
 PT New composition comprising a substance that modulates the status of
 PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
 PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
 PT lung cancer.
 XX
 PS Example 2; SEQ ID NO 24; 211pp; English.
 XX
 CC The invention relates to a composition comprising a substance that
 CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed
 CC gene) or a molecule that is modulated by 121P1F1 where status of the cell
 CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical
 CC composition comprising the novel composition in a human unit dose form, a
 CC recombinant protein comprising an antigen-binding region of a monoclonal

CC antibody, a non-human transgenic animal that produces an antibody, a
 CC hybridoma that produces an antibody, a single chain monoclonal antibody
 CC that immunospecifically binds to a 121P1F1-related protein (comprising
 CC the variable domains of the heavy and light chains of a monoclonal
 CC antibody), a vector comprising a polynucleotide that encodes a single
 CC chain monoclonal antibody, a polynucleotide that encodes an analogue
 CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
 CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses
 CC 121P1F1 (comprising administering to the cells the composition), treating
 CC a patient who bears cancer cells that express 121P1F1, generating a
 CC mammalian immune response directed to 121P1F1, inducing an immune
 CC response, monitoring 121P1F1 gene products in a biological sample from a
 CC patient who has or who is suspected of having cancer, monitoring the
 CC presence of cancer in an individual and an assay for detecting the
 CC presence of a 121P1F1-related protein or polynucleotide in a biological
 CC sample from a patient who has or who is suspected of having cancer. The
 CC composition may comprise a polynucleotide that comprises a 121P1F1-
 CC related protein coding sequence provided that the coding sequence does
 CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
 CC substance also comprises a polynucleotide that encodes at least one
 CC peptide given in 16 Tables (given in the specification), the peptides
 CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its
 CC splice variants. The composition is useful for detecting, treating or
 CC preventing cancer, preferably prostate cancer, bladder cancer, kidney
 CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
 CC cervical cancer or stomach cancer. The composition can also be used as a
 CC vaccine to treat or prevent cancer that expresses or overexpresses
 CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
 CC sequence is a 121P1F1 protein (full-length or fragment).
 XX
 SQ Sequence 198 AA;

Query Match 96.6%; Score 1011; DB 8; Length 198;
 Best Local Similarity 100.0%; Pred. No. 3.1e-85;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KGLSAEKKRTMMELFSTKQVFOFKDLKIAPKKGTITAMSVKVLQSLVDDGAVDCER 64
 DB 1 KGLSAEKKRTMMELFSTKQVFOFKDLKIAPKKGTITAMSVKVLQSLVDDGAVDCER 60
 QY 65 IGTSNYWAFPSKALHARKHKLKLEVESQSLSGSQGHASLQKSIKAKIGRCETERTRIA 124
 DB 61 IGTSNYWAFPSKALHARKHKLKLEVESQSLSGSQGHASLQKSIKAKIGRCETERTRIA 120
 QY 125 KELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKYAKKXANRWNTDNIPAIKSAKRF 184
 DB 121 KELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKYAKKXANRWNTDNIPAIKSAKRF 180
 QY 185 GFEENKIDRTFGIPEDFD 202
 DB 181 GFEENKIDRTFGIPEDFD 198

RESULT 13
 AAM40043 standard; protein; 190 AA.
 XX
 AC AAM40043;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 3188.
 XX
 KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.

```
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000MO-US034263.
PF
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00486725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH,
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59199.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 4; SEQ ID NO 3188; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA438642-AA442213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 190 AA;
XX
Query Match 93.1%; Score 975; DB 4; Length 190;
Best Local Similarity 100.0%; Pred. No. 6.3e-82;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 16 MMEIFSETKDVPOLKDLKIAPEKEGITTAMSVKEVLQSLVDGMDVCERIGTSNYWAPF 75
DB 1 MMEIFSETKDVPOLKDLKIAPEKEGITTAMSVKEVLQSLVDGMDVCERIGTSNYWAPF 60
XX
QY 76 SKALHARKHKLLEVLSEQLSEGSQKHAISLQKSIKAKIGRCETERTRLAKELSSLRDQRE 135
DB 61 SKALHARKHKLLEVLSEQLSEGSQKHAISLQKSIKAKIGRCETERTRLAKELSSLRDQRE 120
XX
QY 136 QLKAEVEKTKDCDPQVEEIRQANKVAKAANRWTDNIPAIKSMARKKGFPEENKIDRPF 195
DB 121 QLKAEVEKTKDCDPQVEEIRQANKVAKAANRWTDNIPAIKSMARKKGFPEENKIDRPF 180
XX
QY 196 GIPEDFDYID 205
DB 181 GIPEDFDYID 190
XX
RESULT 14
ADD84547
ID ADD84547 standard; protein; 190 AA.
XX
AC ADD84547;
```

```
XX 29-JAN-2004 (first entry)
DT
XX
XX 121PIF1 variant 4 protein.
DE
XX
XX 121PIF1; 121PIF1 modulation; human; chromosome 4q; cytostatic;
XX gene therapy; vaccine; cancer; immune response; immunisation.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO200295009-A2.
XX
XX 28-NOV-2002.
XX
XX 28-FEB-2002; 2002MO-US006242.
XX
XX 05-MAR-2001; 2001US-00799250.
XX (AGEN-) AGENSYS INC.
XX
XX Chailita-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,
XX Jakobovits A;
XX
XX WPI; 2003-156757/15.
XX DR N-PSDB; ADD84546.
XX
XX Composition comprising a substance that modulates the status of 121PIF1,
XX useful in diagnosing, preventing, prognosticating or treating patients
XX with cancer that expresses 121PIF1, such as breast, colon, ovarian or
XX lung cancer.
XX
XX Claim 19; Fig 2F; 285bp; English.
XX
CC The present invention describes a composition (I) comprising a substance
CC that modulates the status of 121PIF1 (gene and encoded protein), or a
CC molecule that is modulated by 121PIF1, where the status of a cell that
CC expresses 121PIF1 is modulated. The human 121PIF1 gene maps to chromosome
CC 4q. (I) has cytostatic activity, and can be used in gene therapy, and in
CC vaccines. The composition (I) can be used for diagnosing, preventing,
CC prognosticating or treating patients with cancer that expresses 121PIF1,
CC such as breast, colon, ovarian or lung cancer. The 121PIF1 gene or its
CC fragment can be used to elicit a humoral or cellular immune response.
CC 121PIF1 antibodies can be used in active or passive immunisation. 121PIF1
CC polynucleotides are useful as probes and primers for the amplification or
CC detection of 121PIF1 genes, as coding sequences for directing the
CC expression of 121PIF1 polypeptides, or as tools for modulating or
CC inhibiting the expression of 121PIF1 genes. The present sequence is used
CC in the exemplification of the present invention.
XX
XX Sequence 190 AA;
XX
Query Match 93.1%; Score 975; DB 7; Length 190;
Best Local Similarity 100.0%; Pred. No. 6.3e-82;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 16 MMEIFSETKDVPOLKDLKIAPEKEGITTAMSVKEVLQSLVDGMDVCERIGTSNYWAPF 75
DB 1 MMEIFSETKDVPOLKDLKIAPEKEGITTAMSVKEVLQSLVDGMDVCERIGTSNYWAPF 60
XX
QY 76 SKALHARKHKLLEVLSEQLSEGSQKHAISLQKSIKAKIGRCETERTRLAKELSSLRDQRE 135
DB 61 SKALHARKHKLLEVLSEQLSEGSQKHAISLQKSIKAKIGRCETERTRLAKELSSLRDQRE 120
XX
QY 136 QLKAEVEKTKDCDPQVEEIRQANKVAKAANRWTDNIPAIKSMARKKGFPEENKIDRPF 195
DB 121 QLKAEVEKTKDCDPQVEEIRQANKVAKAANRWTDNIPAIKSMARKKGFPEENKIDRPF 180
XX
QY 196 GIPEDFDYID 205
DB 181 GIPEDFDYID 190
```

RESULT 15
ADM83857
ID ADM83857 standard; protein; 190 AA.
XX
AC ADM83857;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human cancer gene 121P1F1 protein 16-205.
XX
KM Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
KW cervical cancer; stomach cancer; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN US2003223997-A1.
XX
PD 04-DEC-2003.
XX
PF 28-FEB-2002; 2002US-00087190.
XX
PR 08-FEB-2001; 2001US-00779250.
XX
PA (CHAL/) CHALLITA-EID P M.
PA (HUBE/) HUBERT R S.
PA (RAIT/) RAITANO A B.
PA (PARI/) PARIS M.
PA (AFAR/) AFAR D E H.
PA (GEWM/) GE W.
XX (JAKO/) JAKOBOVITS A.
XX
PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W,
PI Jakobovits A;
XX
DR WPI; 2004-060522/06.
XX
PT New composition comprising a substance that modulates the status of
PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
PT lung cancer.
XX
PS Example 5; SEQ ID NO 67; 211pp; English.
XX
CC The invention relates to a composition comprising a substance that
CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed
CC gene) or a molecule that is modulated by 121P1F1 where status of the cell
CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical
CC composition comprising the novel composition in a human unit dose form, a
CC recombinant protein comprising an antigen-binding region of a monoclonal
CC antibody, a non-human transgenic animal that produces an antibody, a
CC hybridoma that produces an antibody, a single chain monoclonal antibody
CC that immunospecifically binds to a 121P1F1-related protein (comprising
CC the variable domains of the heavy and light chains of a monoclonal
CC antibody), a vector comprising a polynucleotide that encodes a single
CC chain monoclonal antibody, a polynucleotide that encodes an analogue
CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses
CC 121P1F1 (comprising administering to the cells the composition), treating
CC a patient who bears cancer cells that express 121P1F1, generating a
CC mammalian immune response directed to 121P1F1, inducing an immune
CC response, monitoring 121P1F1 gene products in a biological sample from a
CC patient who has or who is suspected of having cancer, monitoring the
CC presence of cancer in an individual and an assay for detecting the
CC presence of a 121P1F1-related protein or polynucleotide in a biological
CC sample from a patient who has or who is suspected of having cancer. The
CC composition may comprise a polynucleotide that comprises a 121P1F1-
CC related protein coding sequence, provided that the coding sequence does
CC not encode the entire amino acid sequence of 121P1F1 (ADM83793). The
CC substance also comprises a polynucleotide that encodes at least one
CC peptide given in 16 Tables (given in the specification), the peptides
CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its

CC splice variants. The composition is useful for detecting, treating or
CC preventing cancer, preferably prostate cancer, bladder cancer, kidney
CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
CC cervical cancer or stomach cancer. The composition can also be used as a
CC vaccine to treat or prevent cancer that expresses or overexpresses
CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
CC sequence is a 121P1F1 protein (full-length or fragment).
XX
SQ Sequence 190 AA;
Query Match 93.1%; Score 975; DB 8; Length 190;
Best Local Similarity 100.0%; Pred. No. 6.3e-82;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 MMEIFSEKTDVYFQKDLKELIAPKEKGTAMSVKVEVLQSLVDGWDVCERIGTSNYMAFP 75
DB 1 MMEIFSEKTDVYFQKDLKELIAPKEKGTAMSVKVEVLQSLVDGWDVCERIGTSNYMAFP 60
QY 76 SKALHARKHKLVELESQSLSEGSQKHASLQKSIKAKIGRCETERTRLAKELSLRDQRE 135
DB 61 SKALHARKHKLVELESQSLSEGSQKHASLQKSIKAKIGRCETERTRLAKELSLRDQRE 120
QY 136 QLKAEVEKXKCDPQVVEBIRQANKVAKAANRWTDNIFALKSMKRRGFENKIDRTF 195
DB 121 QLKAEVEKXKCDPQVVEBIRQANKVAKAANRWTDNIFALKSMKRRGFENKIDRTF 180
QY 196 GIPEDFDYID 205
DB 181 GIPEDFDYID 190

Search completed: July 27, 2005, 09:37:57
Job time : 166 secs

This Page Blank (uspto)

~~UNAVAILABLE COPY~~

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 09:28:34 ; Search time 40 Seconds

(without alignments)
493.111 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047

Sequence: 1 MSKKKGLSNEKRTMWEIF.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir79:*
2: pir1:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	33.3	210	2 T37610	hypothetical coile
2	197	18.8	128	2 T08972	hypothetical prote
3	168.5	16.1	174	2 S61134	hypothetical prote
4	135.5	12.9	1281	2 JCS368	dynactin 1 - mouse
5	130.5	12.5	1053	2 A41642	dynactin 1 - chicken
6	121.5	11.6	1356	2 S32763	kinectin 1 - human
7	117	11.2	768	2 T02572	hypothetical prote
8	116.5	11.1	2442	2 T08621	centrosome associa
9	115.5	11.0	880	2 F75103	conserved hypochet
10	114	10.9	284	2 S24972	tropomyosin alpha,
11	114	10.9	764	2 T05409	hypothetical prote
12	113	10.8	199	2 A32183	tropomyosin TPM1 -
13	112	10.7	284	2 UC6199	alpha-tropomyosin
14	112	10.7	285	2 A24199	tropomyosin NM, sk
15	112	10.7	559	2 S49143	BG10 protein - tap
16	112	10.7	1938	2 A59293	skeletal myosin he
17	111	10.6	559	2 A45620	cyto villin homolog
18	111	10.6	1937	2 I38855	myosin heavy chain
19	110.5	10.6	308	2 T08796	tropomyosin - huma
20	110.5	10.6	629	2 T44607	hypothetical prote
21	110.5	10.6	879	2 C71083	conserved hypochet
22	110	10.5	284	2 I51731	alpha-tropomyosin
23	109.5	10.5	168	2 G86578	CT670 hypothetical
24	109.5	10.5	168	2 B72046	conserved hypochet
25	109.5	10.5	284	2 UC6198	alpha-tropomyosin
26	109.5	10.5	400	2 E70318	hypothetical prote
27	109.5	10.5	1298	2 T24480	hypothetical prote
28	109.5	10.5	1390	2 S51364	sperm tail-specific
29	109	10.4	284	2 JC2551	tropomyosin alpha

30	109	10.4	284	2 S19691	tropomyosin alpha,
31	109	10.4	676	2 S00084	myosin heavy chain
32	108	10.3	670	2 F84899	hypothetical prote
33	107.5	10.3	1085	2 F96712	hypothetical prote
34	107.5	10.3	1137	2 T19414	hypothetical prote
35	107.5	10.3	1169	2 A64505	p115 homolog - Met
36	107	10.2	280	2 A22165	tropomyosin alpha
37	107	10.2	281	2 A34787	tropomyosin 1 alph
38	107	10.2	284	1 TWRBA	tropomyosin alpha
39	107	10.2	284	2 A39816	tropomyosin 2, fib
40	107	10.2	284	2 B27407	tropomyosin alpha
41	107	10.2	284	2 A25825	tropomyosin alpha
42	107	10.2	284	2 A60597	tropomyosin 2, fib
43	107	10.2	955	2 S24348	myosin heavy chain
44	106.5	10.2	746	2 T47237	myosin II heavy ch
45	106	10.1	1558	2 B71603	REGA-H3 antigen PF

ALIGNMENTS

```
RESULT 1
T37610
hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
C1:Species: Schizosaccharomyces pombe
C1:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R1: Hunt, S.; Devlin, K.; Churcher, C.M.; Barrett, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, September 1995
A1:Reference number: Z21730
A1:Accession: T37610
A1:Status: preliminary; translated from GB/EMBL/DBJ
A1:Molecule type: DNA
A1:Residues: 1-210 <HUN>
A1:Cross-references: UNIPROT:Q09739; EMBL:Z54096; PIR:CAA90804.1; GSPDB:GN00066; SPDB:SP
A1:Experimental source: strain 972h-; cosmid G13A11
A1:Gene: SPDB:SPAC13A11.03
A1:Map position: 1
A1:introns: 22/3

Query Match          33.3%; Score 349; DB 2; Length 210;
Best Local Similarity 41.6%; Pred. No. 3.9e-16;
Matches 84; Conservative 36; Mismatches 76; Indels 6; Gaps 5;

QY      5 KGLSAEKRTRMEIFSETKDYFQKDEKAPKKGITANSYKVLQSLVDGKVDGER 64
      ||||| ||||| : || : || ||||| : || : ||||| ||||| : || :
DB      4 KGLSLAEKRTRLEAIFHDSKDFQKVEKLGSK-KQIVLQTVKDVLSLVDNIVKTEK 62
      ||||| ||||| : || : ||||| : || : ||||| ||||| : || :
QY      65 IGTSNVYVAFPEKALHARKHKLVELESQSSQGHASLQKSI--EKAKIGRCETE-EKT 121
      ||||| ||||| : || : ||||| : || : ||||| ||||| : || :
DB      63 IGTSNVYVAFPEKALHARKHKLVELESQSSQGHASLQKSI--EKAKIGRCETE-EKT 122
      ||||| ||||| : || : ||||| : || : ||||| ||||| : || :
QY      122 RLAKELSLRD-QREQLAEVEKYKDCPOVEIRQANKYAKAEANWTDNIFAIKSWA 180
      : || : || : || : || : || : || : || : || : || : || : || :
DB      123 QYTLLELLAKSEELKLTQLSLNHNCHNPFPELKNENTKKYMEANMTQDIIHLIAFC 182
      : || : || : || : || : || : || : || : || : || : || : || :
QY      181 KKKFGFENKIDRTFGIPEDPD 202
      : || : || : || : || : || : || : || : || : || : || : || :
DB      183 RDWGDATNQIREFYCSIPEDLD 203
      : || : || : || : || : || : || : || : || : || : || : || :

RESULT 2
T08972
hypothetical protein F19B15.200 - Arabidopsis thaliana
C1:Species: Arabidopsis thaliana (mouse-ear cress)
C1:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
A1:Accession: T08972
R1: Bevan, M.; Hilbert, H.; Brun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, May 1999
A1:Reference number: Z16519
A1:Accession: T08972
A1:Molecule type: DNA
```


1821
AIOOEOOOOCOFERVKADAIQALFOAHMTJKEPHGELODH---KEDABP 1870

Db 1821 ALQEQQQQAGQGEERYKEKADALQALQAHMTLKERHGELODH----KEQARR 1870

submitted to the Protein Sequence Database, April 1996

A:Reference number: S62997

A:Accession: S63011

A:Molecule type: DNA

A:Residues: 1-199 <POW>

A:Cross-references: EMBL:Z71355; NID:g1301970; PIDN:CAA95953.1; PID:g1301971; MIPS:YNL07

A:Experimental source: strain S288C

R:Solier-Mira, A.; Saliz, J.E.; Ballestra, J.P.G.; Remacha, M.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63018

A:Accession: S63018

A:Molecule type: DNA

A:Residues: 1-199 <SOL>

A:Cross-references: EMBL:Z71355; NID:g1301970; PIDN:CAA95953.1; PID:g1301971; MIPS:YNL07

A:Experimental source: strain S288C

R:Poehlmann, R.; Philippson, P.

Yeast 12, 391-402, 1996

A:Title: Sequencing a cosmid clone of *Saccharomyces cerevisiae* chromosome XIV reveals 12

A:Reference number: S63925; PMID:8701611

A:Accession: S63928

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-199 <POF>

A:Cross-references: EMBL:X86470; NID:g791101; PIDN:CAA60179.1; PID:g791105

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

C:Genetics:

A:Gene: SGD:TPM1

A:Cross-references: SGD:S0005023; MIPS:YNL079C

A:Map position: 14L

C:Superfamily: tropomyosin TPM1

C:Keywords: coiled coil; cytoskeleton

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;

7 LSAEERKRTMMEIFSETKQVDFOLKDEKTAPEKKEGTTAMSEVQLQSLVDD-----GMV 60

13 LRAESQKEKEHEKKNKLEQ-ENVE-----KENQKSLTVKN--QQLDEDEIKLEAGLS 65

61 DCEIRGTSNYWAFPSKALHARKHKL-----EVLSEQLSEG-----SQK 99

66 DSKQTEQDNVKEKNQKSLTVKNHQLSEELKEALSLASKQSLSEDSHLLQSNNDNFSKX 125

100 HASLQKSIKAKIGKRETERSTR-----LAKELSSLRDQEQKAEV-----KYKCC 147

126 NOQLBEDLEESDPTKLETKETKLESDLKADQLERRVALEEQREWERKNEELTVYEDA 185

148 DPQVVE 153

186 KYELDE 191

Db

Query Match 10.8%; Score 113; DB 2; Length 199;

Best Local Similarity 24.7%; Pred. No. 1;

Matches 46; Conservative 40

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 09:23:03 ; Search time 174 Seconds

(without alignments)
603.312 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047

Sequence: 1 MSKKKGSLSAEKRTIMEIF.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*

2: uniprot_crembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1047	100.0	205	2	Q9BWT6
2	954	91.1	205	2	Q8K396
3	948	90.5	205	2	Q9DDA1
4	644	61.5	220	2	Q6DC61
5	462	44.1	196	2	Q86E28
6	440	42.0	230	2	Q8GYD2
7	427	40.8	207	2	Q6H432
8	349	33.3	210	1	YAS3 SCHPO
9	340.5	32.5	203	2	Q8SUA9
10	340.5	32.5	203	2	Q6WDA3
11	340.5	32.5	203	2	Q7QTX0
12	332.5	31.8	179	2	Q6Q9F9
13	277	26.5	196	2	Q7RH53
14	266.5	25.5	211	2	Q6B0L5
15	221	21.1	225	2	Q7SC55
16	206.5	19.7	225	2	Q6CSX5
17	197	18.8	228	2	Q9S2E5
18	175	16.7	222	2	Q6FL56
19	168.5	16.1	174	1	YGT3 YEAST
20	153.5	14.7	101	2	Q6R679
21	136.5	13.0	910	2	Q7T2F8
22	135.5	12.9	1281	1	DYNA MOUSE
23	134.5	12.8	890	2	Q6AMB1
24	134.5	12.8	890	2	Q6AMB3
25	134.5	12.8	1139	2	Q6I037
26	134.5	12.8	1264	2	Q6NZM3
27	134.5	12.8	1278	1	DYNA HUMAN
28	134.5	12.8	1278	2	Q6MZ23
29	131.5	12.6	1280	1	DYNA RAT
30	130.5	12.5	1224	1	DYNA CHICK
31	126.5	12.1	1232	2	Q6PCJ1

32	126	12.0	609	2	Q8TXA4
33	122.5	11.7	797	2	Q6GLB3
34	122	11.7	10578	2	Q8ISF5
35	122	11.7	18519	2	Q8ISF6
36	122	11.7	18534	2	Q8ISF7
37	121.5	11.6	995	2	Q6FIN4
38	121.5	11.6	1357	1	KTNI HUMAN
39	121.5	11.6	1364	1	KTNI CHICK
40	120.5	11.5	172	2	Q82AV3
41	120.5	11.5	448	2	Q84H59
42	120.5	11.5	1732	2	Q9VJ35
43	120	11.5	882	1	RAS0 PYRPU
44	120	11.5	1120	1	ERCI MOUSE
45	118.5	11.3	1177	2	Q877I1

ALIGNMENTS

RESULT 1

Q9BWT6 PRELIMINARY; PRT; 205 AA.

AC Q9BWT6
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 26, Last annotation update)

DE GAI.
GN Name=GAI;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Solis G., Hoter H.W.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bray N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hogg L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci R., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.B., Scherch A., Schein J.E.,
RA Jones S.J., Maitra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strusberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028916; AAK26168.1; -
DR EMBL; BC032142; AAK2142.1; -
DR InterPro; IPR005647; MndI.
DR InterPro; IPR009058; wing_hlx_DNA_bnd.
DR Pfam; PF03962; MndI; 1.
SQ SEQUENCE 205 AA; 23753 MW; 95B0B14068DA0B51 CRC64;

Query Match 100.0%; Score 1047; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSABEKRTRMEIFSETKDYFQOLKDLKIAPEKGTAMSVKEVLSLVDDGV 60
DB 1 MSKKKGLSABEKRTRMEIFSETKDYFQOLKDLKIAPEKGTAMSVKEVLSLVDDGV 60

QY 61 DCERICTSNYYMAFPSPKALHARKHKLTVLESQSLSEGSQKIASLQSIKAKIGRCETEER 120
DB 61 DCERICTSNYYMAFPSPKALHARKHKLTVLESQSLSEGSQKIASLQSIKAKIGRCETEER 120

QY 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWA 180
DB 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWA 180

QY 181 KRKFGEENKIDRTFGIPEDPDYID 205
DB 181 KRKFGEENKIDRTFGIPEDPDYID 205

RESULT 2
Q8K396 PRELIMINARY; PRT; 205 AA.
AC Q8K396;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE GAW protein.
GN Name=2610034E18Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marinsina K., Parner A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield J.S.,
RA Krzywiński M.I., Skalska U., Smalins D.E., Schmechel A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary tumor;
RA Strausberg R.L.; TISSUE=Mammary tumor;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027741; AAH27741.1;
DR MGD; MGI:1924165; 2610034E18Rik.
DR InterPro; IPRO005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 205 AA; 23849 MW; 122C3FA9E425120 CRC64;

Query Match 91.4%; Score 954; DB 2; Length 205;
Best Local Similarity 89.8%; Pred. No. 2.5e-50;
Matches 184; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

QY 1 MSKKKGLSABEKRTRMEIFSETKDYFQOLKDLKIAPEKGTAMSVKEVLSLVDDGV 60

DB 1 MSKKKGLSABEKRTRMEIFSETKDYFQOLKDLKIAPEKGTAMSVKEVLSLVDDGV 60

QY 61 DCERICTSNYYMAFPSPKALHARKHKLTVLESQSLSEGSQKIASLQSIKAKIGRCETEER 120
DB 61 DCERICTSNYYMAFPSPKALHARKHKLTVLESQSLSEGSQKIASLQSIKAKIGRCETEER 120

QY 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWA 180
DB 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWA 180

QY 181 KRKFGEENKIDRTFGIPEDPDYID 205
DB 181 KRKFGEENKIDRTFGIPEDPDYID 205

RESULT 3
Q9D0A1 PRELIMINARY; PRT; 205 AA.
AC Q9D0A1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Mus musculus 10 days embryo whole body cDNA. RIKEN full-length
enriched library, clone:2610034E18 product:GAW homolog.
GN Name=2610034E18Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RT Meth. Enzymol. 303:19-44(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The PANTOM Consortium;
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtration of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RT Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Katsunari T., Tashiro H., Itoh M.,
RA Yamamoto R., Matsuno H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format

```

RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carininci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayashi N., Hiramoto K., Hiraoka T., Hori F.,
RA Imocani K., Ichi Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saio H., Saio R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK011664; BAB27765.1; -
DR MGD; MGI:1924165; 2610034E18R1k.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 205 AA; 23909 MW; 09368E19FA224021 CRC64;

Query Match 90.5%; Score 948; DB 2; Length 205;
Best Local Similarity 89.3%; Pred. No. 5,6e-50;
Matches 183; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSKKKGLSABEKRTTMMEIFSETKDVFOQLDKLKIAPKEKGTITMSVKEVLSLVDDGMV 60
DB 1 MSKKKGLSABEKRTTMMEIFSETKDVFOQLDKLKIAPKEKGTITMSVKEVLSLVDDGMV 60
QY 61 DCRIGTSNYWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTER 120
DB 61 DCRIGTSNYWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTER 120
QY 121 TELAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKYAKKANRWTDNIFAIKSWA 180
DB 121 AMLAKELFSEFDQROQLKAEVEKYRCDPQVVEIRQANKYAKKANRWTDNIFAIKSWA 180
QY 181 KKKFGFEKKIDRTGICPEDFYID 205
DB 181 KKKFGFEKKIDRTGICPEDFYID 205
QY 181 KKKFGFEKKIDRTGICPEDFYID 205
DB 181 KKKFGFEKKIDRTGICPEDFYID 205

RESULT 4
Q6DC61 PRELIMINARY; PRT; 220 AA.
AC Q6DC61;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Zgc:101017 protein.
GN Name=zgc:101017;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RX PubMed:12477932; DOI=10.1073/pnas.242603899;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altchul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Urdin T.B., Toshlyuk S., Carininci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McKernan K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

```

```

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalela U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RA Director MGC Project;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC078223; AAH78223.1; -
DR InterPro; IPR005647; Mnd1.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 220 AA; 25176 MW; 90DEAA69311F4BF7 CRC64;

Query Match 61.5%; Score 644; DB 2; Length 220;
Best Local Similarity 71.8%; Pred. No. 1.4e-31;
Matches 125; Conservative 20; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKKKGLSABEKRTTMMEIFSETKDVFOQLDKLKIAPKEKGTITMSVKEVLSLVDDGMV 60
DB 1 MSKKKGLSABEKRTTMMEIFSETKDVFOQLDKLKIAPKEKGTITMSVKEVLSLVDDGMV 60
QY 61 DCRIGTSNYWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTER 120
DB 61 DCRIGTSNYWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTER 120
QY 61 DCRIGTSNYWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTER 120
DB 61 DCRIGTSNYWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTER 120
QY 121 TRIAEKLSLRDQREQLKAEVEKYKDCDPQVVEIRQANKYAKKANRWTDNIF 174
DB 121 EDLKEKLTALKGQRMKVELEKYOECDAVVEIRQANKYAKKANRWTDNIF 174
QY 121 EDLKEKLTALKGQRMKVELEKYOECDAVVEIRQANKYAKKANRWTDNIF 174
DB 121 EDLKEKLTALKGQRMKVELEKYOECDAVVEIRQANKYAKKANRWTDNIF 174

RESULT 5
Q6E28 PRELIMINARY; PRT; 196 AA.
AC Q6E28;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Clone ZSD1259 mRNA sequence.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_Taxid=6182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22879925; PubMed=12973349; DOI=10.1038/ng1236;
RA Hu W., Yan Q., Shen D.K., Liu F., Zhu Z.D., Song H.D., Xu X.R.,
RA Wang Z.J., Rong Y.P., Zeng L.C., Wu J., Zhang X., Wang J.J., Xu X.N.,
RA Wang S.Y., Fu G., Zhang X.L., Wang Z.Q., Brindley P.J., McManus D.P.,
RA Xue C.L., Feng Z., Chen Z., Han Z.G.;
RT "Evolutionary and biomedical implications of a Schistosoma japonicum
RT complementary DNA resource."
RT Nat. Genet. 35:139-147(2003).
DR EMBL; AY223066; AAP06089.1; -
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 196 AA; 23163 MW; B30F6F08BD7123F0 CRC64;

Query Match 44.1%; Score 462; DB 2; Length 196;
Best Local Similarity 47.4%; Pred. No. 1.2e-20;
Matches 92; Conservative 34; Mismatches 68; Indels 0; Gaps 0;

QY 11 EKRTRMEIFSETKDVFOQLDKLKIAPKEKGTITMSVKEVLSLVDDGMVDCRIGTSNY 70
DB 2 KSRQMDMPFEKDPFOQLDKLKIAPKEKGTITMSVKEVLSLVDDGMVDCRIGTSNY 61
QY 71 YWAFPSKALHARKHKLKLEVLBSQSGHSLQSIKAKIGRCETTERLAKELSSL 130

```

```

Db      62 FMAFSPKAAQKNNNEKVTGDIHFRNQIFKTRSLNELSKRKTEERNRINLET 121
      131 RDQRQLKAWEKYKCDPOVVEIRQANKAEANRWTDNIFAISNAKRFGEENK 190
      122 KILLESLETLQDLQLEHNDPDRLSLQGOQLVLDNSANRWTDNIFVKSMLSNKFSLEANT 181
Qy      191 IDRTFGIPEDFDYI 204
      182 FCRQFEIPENFDYI 195

```

RESULT 6

```

Q6GYD2 ID Q6GYD2 PRELIMINARY; PRT; 230 AA.
AC 08GYD2;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein At4g29170/F19B15_200 (Hypothetical protein
DE At4g29170).
GN Name=At4g29170/F19B15_200; Synonyms=At4g29170;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Euroside 11; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
[1]
RN SEQUENCE FROM N.A.
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hanan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shim P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK117113; BAC42364.1; -.
DR EMBL; BT005435; AA063855.1; -.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
KW Hypothetical protein.
SQ
SEQUENCE 230 AA; 26402 MW; 131BB5146D8C91ED CRC64;

```

Query Match 42.0%; Score 440; DB 2; Length 230;
 Best Local Similarity 43.8%; Pred. No. 3.2e-19;
 Matches 91; Conservative 43; Mismatches 70; Indels 4; Gaps 2;

```

Qy      1 MSKKKGLSAEKKRTMMEIFSETKDVQFLQDLKELIAPKEKGTAMSVKEVLQSLVDGMY 60
      1 MSKKKGLSLEKREKQMLQIFYESQDFLLKELEKMGPK-KGVISQSVKDVQLSLVDDLV 59
Db      61 DCEIRIGTSNYWAFPSKALHARKHLEVLSESGQKASLOKSIKATIGRCETEE 120
      60 AKXKIGISYFWSLPSCAGNQLRVSQKLESDDLQSGSKRIALVLDQCEALKKGRSESEER 119
Qy      121 TRLAKELSLRDQRLKAWEKYKCDPOVVEIRQANKAEANRWTDNIFAISNA 180
      120 TEALTLQKDLKELKHLKEMVQPADNDPATLEKRAIVAIQASANRWTDNIFTLRQNC 179
Qy      181 KRKFGFEENKIDRTF--GIPEDFDYI 205
      180 SNNFPQAKQLEHLTYEAGITDEFDYIE 207
Db

```

RESULT 7

```

Q6H432 ID Q6H432 PRELIMINARY; PRT; 207 AA.
AC Q6H432;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Putative GAI protein.
GN Name=P0651G05.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxId=39947;
[1]
RN SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
RT clone: P0651G05."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006528; BAD26517.1; -.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ
SEQUENCE 207 AA; 23900 MW; 6A9C5ECBEB6D586E CRC64;

```

Query Match 40.8%; Score 427; DB 2; Length 207;
 Best Local Similarity 42.5%; Pred. No. 1.7e-18;
 Matches 88; Conservative 44; Mismatches 71; Indels 4; Gaps 2;

```

Qy      1 MSKKKGLSAEKKRTMMEIFSETKDVQFLQDLKELIAPKEKGTAMSVKEVLQSLVDGMY 60
      1 MSKKKGLSLEKREKQMLQIFYDSQDFLLKELEKMGPK-KGVISQSVKDVQSLVDDLV 59
Db      61 DCEIRIGTSNYWAFPSKALHARKHLEVLSESGQKASLOKSIKATIGRCETEE 120
      60 LKDKIGTSYFWSLPSCAGNQLRTTYSKLESDDLSSKKRFLIYVQRENLRKGRSDSER 119
Qy      121 TRLAKELSLRDQRLKAWEKYKCDPOVVEIRQANKAEANRWTDNIFAISNA 180
      120 EALBELLAWEHKKLEELAAVSDPALAEANNDALIEVAHAANRWTDNIFTLQWC 179
Db      181 KRKFGFEENKID--RTFGIPEDFDYI 204
      180 STTFPQAKQLEHMYREVIGITDEFDYIL 206
Qy
Db

```

RESULT 8

```

YA53 SCHPO ID YA53 SCHPO STANDARD; PRT; 210 AA.
AC Q09739;
DT 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 05-JUL-2004 (rel. 44, Last annotation update)
DE Hypothetical protein C13A11.03 in chromosome 1.
GN ORFNames=SPAC13A11.03;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxId=4896;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris P., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

```

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymporter B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Meller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forstburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpkovski G.V., Ussery D., Barrett B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe.",
RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: 254096; CAA90804.1; -
DR PIR: T37610; T37610
DR GenBank: SPAC13A11.03; -
DR InterPro: IPR005647; Mnd1.
DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
DR Pfam: PF03962; Mnd1; 1.
KW Hypothetical protein.
SQ SEQUENCE 210 AA; 24224 MW; F4A546F070A37665 CRC64;

Query Match 33.3%; Score 349; DB 1; Length 210;
Best Local Similarity 41.6%; Pred. No. 9e-14;
Matches 84; Conservative 36; Mismatches 76; Indels 6; Gaps 5;

QY 5 KGLSAEKRTRMMEIFSETKDFVQLKDLKIAPEKKGITAMSVKYLQSLVDGAVDCER 64
DB 4 KGLSLAEKRRRLAEIAPHDKDFVQLKVEKLGSK-KQIVLQVYKDLQSLVDNIVKTEK 62
QY 65 IGTSNYYAFPSKALHARKHKLVEVLSQSGSKASLQKSI--EKAKIGRCETE-EKT 121
DB 63 IGTSNYYAFPSKAKRSRESVLSQALQDLKQKSKTIDENISFKSKRDNBGTENDAN 122
QY 122 RLAKELSLRD-QREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWA 180
DB 123 QYTLLELHAKESLKLKTLQSLNLHNCNPETPELKNENKRWKWEAANLMTDQHLTLAFC 182
QY 181 KKKFGEENKIDRTFGIPEDFD 202
DB 183 -RDMGADTQIREYCSIPEDLD 203

RESULT 9
Q8SUA9 PRELIMINARY; PRT; 203 AA.
AC Q8SUA9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DN Hypothetical protein ECU10_1600.
GN Name=ECU10_1600;
OS Encephalitozoon cuniculi GB-M1.
OC Encephalitozoon cuniculi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxId=284813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;
RA Katinka M.D., Duprat S., Cornillot E., Metenier F., Thomarat F.,
RA Prensier G., Barde V., Peyretailade E., Brottier P., Wincker P.,

RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vives C.P.,
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.",
RL Nature 414:450-453(2001).

CC SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RA Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL590449; CAD25881.1; -
DR InterPro: IPR002114; HPR_SerP_S.
DR InterPro: IPR005647; Mnd1.
DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
DR Pfam: PF03962; Mnd1; 1.
DR PROSITE: PS00589; Pts_HPR_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 203 AA; 23743 MW; DFE33A65A1A28A42 CRC64;

Query Match 32.5%; Score 340.5; DB 2; Length 203;
Best Local Similarity 35.1%; Pred. No. 2.8e-13;
Matches 71; Conservative 50; Mismatches 72; Indels 9; Gaps 3;

QY 7 LSABEKRTRMMEIFSETKDFVQLKDLKIAPEKKGITAMSVKYLQSLVDGAVDCERIG 66
DB 6 MKSDQKSLILHEIRIGSKSPFKQLQELSLGSK-KGIYVNTIKELIQLVLDGLVTAERKG 64
QY 67 TSNYYAFPSKALHARKHKLVEVLSQSGSKASLQKSI-EKAKIGRCETEERTRLAKE 126
DB 65 TSNLYWSPASBGIOKKLRCKELMEBCERMSQDICRKEVYLENEMSGHYEERELBNK 124
QY 127 LSSL-----RDREQKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWA 182
DB 125 LWALKTIQDQRE----ELGKFEETDPIAYKLVADRKEMADBCNRITIDNVIIDYICS 180
QY 183 KKFGEENKIDRTFGIPEDFDYI 204
DB 181 KFPMEKSEFNSFGIPEDLDYI 202

RESULT 10
Q6MDA3 PRELIMINARY; PRT; 203 AA.
AC Q6MDA3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DB Mnd1.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX NCBI_TaxId=57411;
RN [1]
RP SEQUENCE FROM N.A.
RA Ramesh M.A., Malik S.B., Logsdon J.M. Jr.,
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY295092; AAQ24512.1; -
DR InterPro: IPR005647; Mnd1.
DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
DR Pfam: PF03962; Mnd1; 1.
SQ SEQUENCE 203 AA; 23278 MW; CC0625DCDC4158C0 CRC64;

Query Match 32.5%; Score 340.5; DB 2; Length 203;
Best Local Similarity 34.8%; Pred. No. 2.8e-13;
Matches 70; Conservative 46; Mismatches 84; Indels 1; Gaps 1;

QY 5 KGLSAEKRTRMMEIFSETKDFVQLKDLKIAPEKKGITAMSVKYLQSLVDGAVDCER 64
DB 4 KGLSLDEKKERLEBMLKRGELYSNKTLETLS-KPTGISSWIKVLAQVAVEDLVTDK 62
QY 65 IGTSNYYAFPSKALHARKHKLVEVLSQSGSKASLQKSI-EKAKIGRCETEERTRLA 124
DB 63 IGASTYYWCFASKRSQAARTELARQLAEQTFNFIIDATARIELKVGRETERSSLL 122

```
QY 125 KELLSSLRDQEQKAEVEKTKDCDPQVEIRQANKVAKAANRWTDNIFAISMAKRF 184
DB 123 KELLALQVLEEQRGTFRDILKNDDPVAQKLRYTIDIAKQEANLWTDNIFCLOKMYLTKL 182
QY 185 GFEENKIDRTFGJPEDEFDYD 205
DB 183 QMDKKTVSTALGITGEFDYLE 203

RESULT 11
Q7OTX0 PRELIMINARY; PRT; 203 AA.
AC Q7OTX0;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE GIP 76 12374 11763.
OC Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxId=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AACB0100100; EAA38463.1; -.
DR InterPro: IPR005647; Mnd1.
DR InterPro: IPR009058; WING_hlx_DNA_bnd.
DR Pfam: PF03962; Mnd1.
SQ SEQUENCE 203 AA; 23278 MW; CC0625DCC4158C0 CRC64;

Query Match
Best Local Similarity 32.5%; Score 340.5; DB 2; Length 203;
Matches 70; Conservative 46; Mismatches 84; Indels 1; Gaps 1;

QY 5 KGLSAEKRTRMEIFSETQDVOLFQDLKLEIAPEKGITAMSVKEVLSVDDGMDCR 64
DB 4 KGTSLDEKKEKRLLEMLKRGELISNKTITLTS-KPTGISSMVKLVQALVNDVDVDTK 62
QY 65 IGTSNYYWAFPSFALHARKHKLVLVSQSEGSQKASLQKSIKAKIGCETEEETRLA 124
DB 63 IGASTYYWCFASGRSQAAARTELARLQKALEEQTNFIDKATARIIEELKVGRETEERSILL 122
QY 125 KELLSSLRDQEQKAEVEKTKDCDPQVEIRQANKVAKAANRWTDNIFAISMAKRF 184
DB 123 KELLALQVLEEQRGTFRDILKNDDPVAQKLRYTIDIAKQEANLWTDNIFCLOKMYLTKL 182
QY 185 GFEENKIDRTFGJPEDEFDYD 205
DB 183 QMDKKTVSTALGITGEFDYLE 203

RESULT 12
Q6O9F9 PRELIMINARY; PRT; 179 AA.
AC Q6O9F9;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE NMD1 domain containing protein.
OC Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
OX NCBI_TaxId=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Chandra P.K., Wikel S.K.;
```

```
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY552061; AAS79347.1; -.
DR InterPro: IPR005647; Mnd1.
DR Pfam: PF03962; Mnd1.
SQ SEQUENCE 179 AA; 20900 MW; 6E0016920350E476 CRC64;

Query Match
Best Local Similarity 31.8%; Score 332.5; DB 2; Length 179;
Matches 66; Conservative 52; Mismatches 57; Indels 3; Gaps 2;

QY 1 MSK-KKGLSAEKRTRMEIFSETQDVOLFQDLKLEIAPEKGITAMSVKEVLSVDDGM 59
DB 1 MSK-KKGLSAEKRTRMEIFSETQDVOLFQDLKLEIAPEKGITAMSVKEVLSVDDGM 59
QY 60 VDCERIGTSNYYWAFPSFALHARKHKLVLVSQSEGSQKASLQKSIKAKIGCETEE 119
DB 61 VETDKIGSQYYWFPFAPKSRKXKQVLFQDLKLEIAPEKGITAMSVKEVLSVDDGM 120
QY 120 KELLSSLRDQEQKAEVEKTKDCDPQVEIRQANKVAKAANRWTDNIFA 175
DB 121 SSEMEFKLNTLKEKQKSLSKLAKKQSDQNSVDKNNRNLPLDHDANRWDTTYS 178

RESULT 13
Q7RH53 PRELIMINARY; PRT; 196 AA.
AC Q7RH53;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Homo sapiens GAD, putative.
GN Name=PY04140;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertia M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmiller S.B., Feldlyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoah A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Valdivia A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519 (2002).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AABL01001237; EAA15959.1; -.
DR InterPro: IPR005647; Mnd1.
DR Pfam: PF03962; Mnd1.
SQ SEQUENCE 196 AA; 23066 MW; 7FD8820707329074 CRC64;

Query Match
Best Local Similarity 26.5%; Score 277; DB 2; Length 196;
Matches 60; Conservative 52; Mismatches 76; Indels 14; Gaps 3;

QY 3 KKKGLSAEKRTRMEIFSETQDVOLFQDLKLEIAPEKGITAMSVKEVLSVDDGMDC 62
DB 2 KKKGSNEDKILIVDIMESEFFILKELELAAR-KGINSIFKDLIQULIDNKTKS 60
QY 63 ERTGSNYYWAFPSFALHARKHKLVLVSQSEGSQKASLQKSIKAKIGCETEE 115
DB 61 EKVGSQNVFWILKTESSILQNKYQELDKDKREYEMQAKEKNYAELENSLS-----L 114
QY 116 ETEERTLAKELSSLRDQEQKAEVEKTKDCDPQVEIRQANKVAKAANRWTDNIFA 175
```


Db 115 KTDELKOTLKEVKYVLDISEIKKSELDKLTIDIKQIEKMKIQSNFATESIERMANNIFL 174
 QY 176 IKSNAKKKFGFEENKIDRTFGI 197
 Db 175 LKQWTDRTKNSGVDVRLGLM 196

RESULT 14

Q6B0L5 PRELIMINARY; PRT; 211 AA.
 ID Q6B0L5
 AC Q6B0L5;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Similar to CA3413|IPF9239 Candida albicans IPF9239 unknown function (Fragment).
 GN ORFNames=DEHA0E04994g;
 OS Debaryomyces hansenii CBS767.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
 NCBI_Taxid=284592;
 RN [1]
 SQ SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RG Genolevures;

RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Latouche I., de Montigny J., Marc C., Neveglisse C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barde V.,
 RA Barry S., Blanchin S., Beckerich J.M., Beyne E., Bleykaesten C.,
 RA Boissane A., Boyer J., Catolico L., Confantoleri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hameguin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Pellenc J.M., Nikolski M., Oztas S., Olier-Kalogeropoulos O.,
 RA Sennene D., Peltier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swenne D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Wincker C., Caudon B., Scarpelli C., Gaillardin C., Welschenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts."
 RL Nature 430:35-44(2004).
 RN [2]
 SQ SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RG Genoscope;
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: CR382137; CAG87731.1; --
 DR InterPro: IPR005647; Mnd1.
 DR Pfam: PF03962; Mnd1; 1.
 FT NON_TER 1
 SQ SEQUENCE 211 AA; 24831 MW; F8B5D6C6618334C8 CRC64;

Query Match 25.5%; Score 266.5; DB 2; Length 211;
 Best Local Similarity 29.0%; Pred. No. 8.8e-09;
 Matches 61; Conservative 55; Mismatches 75; Indels 19; Gaps 5;

QY 4 KKGLSAEKRTKMEIFSETKDVFLQKLEKIAPKEKGITAMSVKVLQSLVDGMVCE 63
 Db 3 KKGLSAEDKCKELVEFPNQSHFTYLLKEIKESGKYAKISSMLKIDIVQQLIDMNLNCE 62
 QY 64 RIGTSNYTVAFPSSKALHARKKLEVLVESQLS--EGSQKASLQKS--IEKATIGRCET-- 117
 Db 63 KGGTTLNLYWCF-----KFDKIKTLQTYNNYQNKLEKQKLERDQLIEKTLQGLQRLV 115
 QY 118 ----BERTLAKELSLRDQREOLKAVEKYKDCDPQVEERIRQANKVAKAANRWTDNI 173
 Db 116 KSDPGRNRLIDQFCLSKRKHLBEBELKAGDNDPQLIQITDEKRVHLLALETFTDDI 175
 QY 174 FAIKSMARK--KFGFEENKIDRTFGIPEDF 201
 Db 176 ESMIYFTKVSATIEELDLRSELGIPSEF 205

RESULT 15

Q75CH5

PRELIMINARY; PRT; 225 AA.

ID Q75CH5
 AC Q75CH5;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE ACL056CD.
 GN ORFNames=ACL056C;
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Eremothecium.
 NCBI_Taxid=33169;
 RN [1]
 SQ SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10895;
 RA Brachat S., Voegelé S.E., Dietrich F.S., Lerch A., Gaffney T.,
 RA Philippson P.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE016816; AAS51172.1; --
 DR AGD; ACL056C; --
 DR InterPro: IPR005647; Mnd1.
 DR Pfam: PF03962; Mnd1; 1.
 SQ SEQUENCE 225 AA; 25753 MW; BA73C889FFB8B90F CRC64;

Query Match 21.1%; Score 221; DB 2; Length 225;
 Best Local Similarity 26.4%; Pred. No. 5.3e-06;
 Matches 56; Conservative 51; Mismatches 93; Indels 12; Gaps 5;

QY 3 KKGLSAEKRTKMEIFSETKDVFLQKLEKIAPKE-KGITAMSVKVLQSLVD-DGMV 60
 Db 4 KRAVYTLAEKARVILKFGQEHSHSTYSIDLEKLPKKAGVSMVLYQQLIDEDDLI 63
 QY 61 DCERIGTSNYTVAFPSSKALHARKKLEVLVESQLSEGSQKASLQKSIE-----KAKIGR 114
 Db 64 SVEKGGANNVWVCFNQVLGVMKCTEMQMKARSEBSQVRLQDLQAINSEKKHARAALFR 123
 QY 115 CETBERT--LAKELSSLRDQREOLKAVEKYKDC--DPQVEERIRQANKVAKAANRW 170
 Db 124 SEGVSYTRQALLTEHDELGRQLAALQSAVRYRLDPTKMDETKIDSYCRGVRSKLEQLDKIT 183
 QY 171 DNFAIKSMARKKFGFEENKIDRTFGIPEDDP 202
 Db 184 DNEIVYSFLMRRAVSRALAAALDMPERF 215

Search completed: July 27, 2005, 09:40:57
 Job time : 176 secs

~~BEST AVAILABLE COPY~~

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 09:29:49 ; Search time 42 Seconds
(without alignments)
364.359 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047
Sequence: 1 MSKKKGLSABEKRTIMEIF.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	593.5	56.7	127	4	US-09-621-976-4959 Sequence 4959, Ap
2	191.5	18.3	260	4	US-09-248-796A-14485 Sequence 14485, A
3	134.5	12.8	1270	4	US-09-538-092-1321 Sequence 1321, Ap
4	117	11.2	976	3	US-09-104-324B-4 Sequence 4, Appl
5	117	11.2	976	4	US-09-538-092-1339 Sequence 1339, Ap
6	114	10.9	284	4	US-09-914-259-55 Sequence 55, Appl
7	113.5	10.7	284	4	US-09-914-259-62 Sequence 62, Appl
8	112.5	10.8	630	4	US-09-248-796A-20275 Sequence 20275, A
9	112	10.7	284	4	US-09-914-259-43 Sequence 43, Appl
10	111	10.6	1337	4	US-09-538-092-918 Sequence 918, App
11	110	10.5	284	4	US-09-914-259-46 Sequence 46, Appl
12	110	10.5	284	4	US-09-914-259-49 Sequence 49, Appl
13	109.5	10.5	170	4	US-09-438-185A-708 Sequence 708, Appl
14	109	10.4	284	4	US-09-914-259-51 Sequence 51, Appl
15	108	10.3	817	4	US-09-248-796A-20276 Sequence 20276, A
16	107	10.2	281	4	US-09-914-259-63 Sequence 63, Appl
17	107	10.2	284	4	US-09-167-206-10 Sequence 10, Appl
18	107	10.2	284	4	US-09-914-259-40 Sequence 40, Appl
19	107	10.2	284	4	US-09-914-259-48 Sequence 48, Appl
20	107	10.2	284	4	US-09-914-259-50 Sequence 50, Appl
21	106.5	10.2	284	4	US-09-914-259-41 Sequence 41, Appl
22	106	10.1	534	4	US-09-103-664A-2 Sequence 2, Appl
23	106	10.1	1786	3	US-08-973-462-8 Sequence 8, Appl
24	105.5	10.1	224	2	US-08-272-255-16 Sequence 16, Appl
25	105.5	10.1	224	5	PCT-US95-08565-16 Sequence 16, Appl
26	105.5	10.1	245	4	US-10-164-595-34 Sequence 34, Appl
27	105	10.0	245	4	US-09-914-259-65 Sequence 65, Appl

28	105	10.0	251	4	US-09-914-259-64	Sequence 64, Appl
29	105	10.0	284	4	US-09-914-259-60	Sequence 60, Appl
30	104	9.9	1939	4	US-09-538-092-915	Sequence 915, App
31	104	9.9	1939	4	US-09-949-016-11104	Sequence 11104, A
32	102	9.7	281	4	US-09-914-259-45	Sequence 45, Appl
33	102	9.7	284	4	US-09-914-259-47	Sequence 47, Appl
34	102	9.7	284	4	US-09-914-259-57	Sequence 57, Appl
35	102	9.7	372	1	US-07-813-584A-3	Sequence 3, Appl
36	102	9.7	372	1	US-08-330-515-3	Sequence 3, Appl
37	102	9.7	1031	4	US-09-914-259-24	Sequence 24, Appl
38	102	9.7	1940	4	US-09-538-092-901	Sequence 901, App
39	102	9.7	1963	4	US-09-949-016-8888	Sequence 8888, Ap
40	101.5	9.7	2662	4	US-09-595-684B-31	Sequence 31, Appl
41	101.5	9.7	2663	4	US-09-538-092-1252	Sequence 1252, Ap
42	100.5	9.6	284	4	US-09-914-259-59	Sequence 59, Appl
43	100	9.6	1104	3	US-08-923-992A-4	Sequence 4, Appl
44	100	9.6	1164	3	US-08-923-992A-10	Sequence 10, Appl
45	99.5	9.5	588	4	US-08-714-741-42	Sequence 42, Appl

ALIGNMENTS

```
RESULT 1
US-09-621-976-4959
; Sequence 4959, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Maline Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4959
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4959

Query Match      56.7% Score 593.5; DB 4; Length 127;
Best Local Similarity 89.6%; Pred. No. 1.7e-50;
Matches 120; Conservative 5; Mismatches 2; Indels 7; Gaps 1;

QY      1 MSKKKGLSABEKRTIMEIFSETKDVFLQKLEKIAPEKGTITANSYKVLQSLVDGMV 60
      |||||
DB      1 MSKKKGLSABEKRTIMEIFSETKDVFLQKLEKIAPEKGTITANSYKVLQSLVDGMV 60

QY      61 DCEKIGTSNYWAPFSPKALHARKKLEVLLESQSLSGSOKHSLQSIKAKIGRCETER 120
      |||||
DB      61 DCEKIGTSNYWAPFSPKALHARKKLEVLLESQSLSGSOKHSLQSIKAKIGRCET--- 117

QY      121 TRLAKELSLRDR 134
      |||||
DB      118 ----IKLSGMQER 127

RESULT 2
US-09-248-796A-14485
; Sequence 14485, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.133
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
```

```
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14485
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14485
```

```
Query Match      18.3%; Score 191.5; DB 4; Length 260;
Best Local Similarity 27.1%; Pred. No. 8.8e-11;
Matches 59; Conservative 4; Mismatches 75; Indels 37; Gaps 8;
```

```
QY      1 MSKKGLSAEKKTRMMEIFSETKDVFLKDLKIAPEKGTAMSVKVLQSLVDDGMV 60
DB      60 MPKKGLTQEBKSLALINMFQSDHMFYTLKEISKASKQCKIPPMQKELVLAIVEGLV 119
QY      61 DCRIGTSNYMAFPSPKALHARKHKLVELESQSGKASLOKSEKAK-----IGR 114
DB      120 EDRCGITLYWSPF-----YLQHKQ-----QETHRLNRTIANLETERDSLICK 165
QY      115 C--ETERTFLAKELSLR-----DREQQLKAEVEKYKDCDPQVVEIRQANKVAKAA 166
DB      166 CQETGVNRQTHRAKIRFCDSLSERIDISQQLQSGKDSKES--VE-----NLVTSIAF 218
QY      167 NRTDNIIFAIKSWAKKRFGEENKIDRTFGIPDPDYI 204
DB      219 --FSDSIDDIICYLISROTGLTMTTLKTEFELPLEFERI 254
```

```
RESULT 3
US-09-538-092-1321
; Sequence 1321, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1321
; LENGTH: 1270
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number Q14203
US-09-538-092-1321
```

```
Query Match      12.8%; Score 134.5; DB 4; Length 1270;
Best Local Similarity 23.4%; Pred. No. 0.00028;
Matches 50; Conservative 4; Mismatches 106; Indels 15; Gaps 5;
```

```
QY      2 SKKGLSAEKKTRMMEIFSETKDVFLKDLKIAPEKGTAMSVKVLQSLVDDGMV 61
DB      271 ARKEAKEALEAKERYEMADTDADIAEMATLDEKMAERASLQOEVEALKEKRVDELTTD 330
QY      62 CEKI-----GTSNYMAFPSPKALHAR-KHKLVELESQSGKASLOKSEKAK 110
DB      331 LELIKAEIEKSGDGAASVQLKQLEQNARKLDALVRMDLSSSEKQEHVKLQCKMEK- 389
QY      111 KIGRCET--BERTFLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAAAR 168
DB      390 KNOELVVRQQRERLOEELSQAESTIDELKEQVDALGAE-ENVEMLTDRNLWLEEKVRE 448
```

```
QY      169 WTDNIFAIKSWAKKRFGEENKIDRTFGIPDPD 202
DB      449 LRETVGDLEANNEMNDELQENARETELEIREQLD 482
```

```
RESULT 4
US-09-104-324B-4
; Sequence 4, Application US/09104324B
; Patent No. 6232460
; GENERAL INFORMATION:
; APPLICANT: T recit, Ozlem; Sahin, Ugur; Pfreundschuh, Michael
; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,
; TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers In A Sample Of
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,324B
; FILING DATE: 25-June-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/892,702
; FILING DATE: 15-July-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6232460man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5491
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-104-324B-4
```

```
Query Match      11.2%; Score 117; DB 3; Length 976;
Best Local Similarity 23.9%; Pred. No. 0.0099;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;
```

```
QY      3 KKKGLSAE--KRTMMEL--FSETKDV-----FOLKDLKIAPEKGTAMSVKVL 51
DB      413 QKKSSLEBMTGLTNKKEVLEELKKVLGEKETLYENKQFKIABELKG-TEQELIGLL 471
QY      52 QSL---VDDGMVDCERIGTSNYMAFPSPKAL-----HARKKLEYLE-- 90
DB      472 QAREKEVHLEIQVLAITTSBQYSGKVDLKTLENEKDKNTETLSHCNKLSENKEELT 531
QY      91 -----SOLSEGSOKHASLQSIKAKIGRETEFERFLAKELSSLRDQRE 136
DB      532 QETSDMTLELKNQOBDINNKKQOBERMLQIE--NLQETETQLRNLEVEVREELQKDE 589
QY      137 LAAEVEKYKDCD-----POVVEIRQANKVAKS--AANRTDNIFAIK----- 177
DB      590 VCKCKDKSEKCNLRKQVENAKKIIEELQENKALKKGTASQQLNVEIKVANKLEDE 649
QY      178 -SWAKKRFG-----FEENKI 191
DB      650 LSAKQKFGELITDYQKEIEDKKI 673
```

```
RESULT 5
```

```

US-09-538-092-1339
; Sequence 1339, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Lotic
; APPLICANT: Manfred, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1339
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q15431
US-09-538-092-1339

Query Match          11.2%; Score 117; DB 4; Length 976;
Best Local Similarity 23.9%; Pred. No. 0.0099;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;

QY 3 KKKGSAE--KRTMMEI--FSETKOV-----FOLDLEKIAPEKGTAMSVKYL 51
DB 413 QKKSELEBMTKLTKNKEVELKLVGCEKSTLYENKQFKIABELKG--TEQELIGL 471
QY 52 OSI---VDDGVDCRIGTSNYWAFPSKAL-----HARKHLEVL-- 90
DB 472 QAREKEVDHLEIQLRATITSEQYSGEVKVDLKTLENEKLTNTLSTHCKSLNKELT 531
QY 91 -----SOLSGSQKASLQKSIKAKIGRCETEERTRLAKELSLDQREQ 136
DB 532 QETSDMTLELKNQOEINNNKQOEERMLKQIE--NIQETETQLRNELEYREELKQKDE 589
QY 137 LKAEEVKYKD-CD-----POVEEIRQANKVAK--AANRWTDNIFAK----- 177
DB 590 VACKLDKSEKCNMLRKQVENKQKTIIELOQENKALKKGKTAESQLANVETIKVKNLEL 649
QY 178 -SWAKRKFQ-----FEENKI 191
DB 650 LESAKQKGEITDYOKEIEDKKI 673

RESULT 6
US-09-914-259-55
; Sequence 55, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-914-259-55

Query Match          10.9%; Score 114; DB 4; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.0037;

```

```

Matches 52; Conservative 37; Mismatches 61; Indels 74; Gaps 11;

QY 1 MSKKGSAE---EKRTMMEIFSETKOVFQDKLEKIAPEKGTAMSVKESLQSLVD 57
DB 73 LAERKADDAADVASLNRRIQLFEEBLDRAQ---ERLA-----TALQKEAEKAD 122
QY 58 ---GVDCERIGTSNYWAFPSKALHAR---KHLVLESQSLSGSKH----- 100
DB 123 SERGM-----KYIESRAQDEEMELQELQKKA--KHIAEDADRYE 163
QY 101 -----ASLQKSIKAKI--GRCTEERTRLAKELSLDQREQDIAVEKYKCD 148
DB 164 EVARKLVIEESDLAEAEERAEISGKC---AEIEBELKVTYNNLKSLEAQERYSQKE 218
QY 149 POVEEIRQANKVAKAANRWTDNIFAKSMKRFGEENKID 192
DB 219 DKYEERIKVLSDKLKEATR-----AEFAERSVTLEKSID 254

RESULT 7
US-09-914-259-62
; Sequence 62, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-914-259-62

Query Match          10.8%; Score 113.5; DB 4; Length 284;
Best Local Similarity 21.1%; Pred. No. 0.0042;
Matches 48; Conservative 48; Mismatches 78; Indels 53; Gaps 8;

QY 2 SKKKLSA--EKRTMMEIFSETKOVFQDKLEKIAPEKGTAMSVK--EVLQSLVDGM 59
DB 45 AKKLLRASDEDRVLELHKAEDSLAAD-ETAKKADADVASLNRRIQLVEEBLDR 103
QY 60 VDCERIGTSNYWAFPSKALHARKHLEVLSEKSGSK----- 99
DB 104 ---ERLATALQKLEAEKADAESEKGMKVIISRAQKDEKVEIQEIQLEKAKHIAEDADR 160
QY 100 -----HASLQKSIKAKI--GRCTEERTRLAKELSLDQREQDIAVEKYK 145
DB 161 KYEEVARKLVIEESDLAEAEERAEISGKC---AEIEBELKVTYNNLKSLEAQERYKS 215
QY 146 DCDPOVEEIRQANKVAKAANRWTDNIFAKSMKRFGEENKID 192
DB 216 QKEDYEERIKVLSDKLKEATR-----AEFAERSVTLEKSID 254

RESULT 8
US-09-248-796A-20275
; Sequence 20275, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13

```

```
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20275
; LENGTH: 630
; TYPE: PRN
; ORGANISM: Candida albicans
US-09-248-796A-20275
```

```
Query Match          10.7%; Score 112.5; DB 4; Length 630;
Best Local Similarity 23.8%; Pred. No. 0.015;
Matches 49; Conservative 43; Mismatches 79; Indels 35; Gaps 8;
```

```
QY      3 KKKGLSAEKKRTMMEIFSETKOVFOLKDIETAPKEGTT--AMSVEKVLQSLVDGMY 60
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      158 KTKNSDTELEKLEKELEKVK-----LDLQTADEKLGITEREIALKSELFTVKN-- 210
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      61 DCEKIGTSTNYMFP--KALHARKHLEVL-----ESQSESGOKXASLOKSI----- 107
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      211 ---LSTSELALTYTVNSLSEKEBELPLSGNKSKELEDYIQKSDISEKULATDEL 266
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      108 -EKAKIGCETERTLAKELSLRDQEQKAEVEKYKCDPQVVEIRQANKVAKEA 166
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      267 KERTKGFDSKSKLTLENDLSTKKLETEKTQTSKFNLEBKDKKEIVKLNKELELK 326
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      167 NRRTDNIFAIKSAKAKKFGFEENKID 192
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      327 N---DN-----SGAKKELEKVKSLB 344
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

```
RESULT 9
US-09-914-259-43
; Sequence 43, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 284
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-914-259-43
```

```
Query Match          10.7%; Score 112; DB 4; Length 284;
Best Local Similarity 23.5%; Pred. No. 0.0059;
Matches 52; Conservative 41; Mismatches 72; Indels 56; Gaps 9;
```

```
QY      1 MSKKGLSABEK--RTRMMEIFSETKOVFQ-----LKDLEKAPK-EKGITAMSV 47
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      73 LAEKKADAEAVASINRRIOVEEBLDRQERLATLOKLEAEKAADESERGMKVLEN 132
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      48 KEVLQSLVDGMYDCEKIGTSTNYMFP--KALHARKHLEVLSESGOKXASLOKSI 107
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      133 R---ALKDEKKELEKLEI-----OLKEAKHIAEADKKEEVAKKVIIEGDL 176
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      108 EKAKIGCETERTLAKELSLRDQEQKAEVEKYKCDPQVVEIRQANKVAKEA 159
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      177 ER-----TEBAELAESKCSLEBELKVTNNLKSLEAQAKEYSGKEDKYEIEIKILT 229
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      160 KVAKEAANRTDNIFAIKSAKAKKFGFEENKIDRTFGIPED 200
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      230 DKLEKAEETRAE---FAERSVA-----KLEKTIIDLD 258
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

```
RESULT 10
US-09-538-092-918
```

```
; Sequence 918, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Iolc
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 918
; LENGTH: 1937
; TYPE: PRN
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P13535
US-09-538-092-918
```

```
Query Match          10.6%; Score 111; DB 4; Length 1937;
Best Local Similarity 23.1%; Pred. No. 0.096;
Matches 54; Conservative 42; Mismatches 88; Indels 50; Gaps 9;
```

```
QY      4 KKKGLSAEKKRTMMEIFSETKOVFOLKDIETAPK-----EKGITAMS-----V 47
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1361 QPALSANSSEVAQMTKTIETDAIQRTHELEAKKLAQRLQAEHVEAVNAKASLEKT 1420
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      48 KEVLQSLVDGMYDCEKIGTSTNYMFP--KALHARKHLEVLSESGOKXASLOKSI 98
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1421 KQRLQNEVEDLMDLVDERSMAACALDKKQRF-----DKVLSEWQKKVETQAELEASQK 1475
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      99 KIASLOKSTIEKAK-----IGRCET--BERTLAKELSLRDQ-----REQKAE 140
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1476 ESRSLSTELFKYKNVYEEISLDLETLRRKNKLOOEISDLOIABGGQIHELEKIKQ 1535
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY      141 VEKYKCDPQVVEIRQANKVAKEA-ANRWNTDNIFAIKSAKAKKFGFEENKIDR 193
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      1536 VEOEK-CEIOALAEAEASLEHBEKILRIQLELQVASEVDKIAEKDEEIDQ 1588
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

```
RESULT 11
US-09-914-259-46
; Sequence 46, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 284
; TYPE: PRN
; ORGANISM: Bradydiano rexio
US-09-914-259-46
```

```
Query Match          10.5%; Score 110; DB 4; Length 284;
Best Local Similarity 23.3%; Pred. No. 0.0092;
Matches 51; Conservative 44; Mismatches 72; Indels 52; Gaps 10;
```

```
QY      1 MSKKGLSABE---EKTRMMEIFSETKOVFQ-----LKDLEKAPK-EKGITAMSV 47
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      73 LAEKKATDAEGDVASINRRIOVEEBLDRQERLATLOKLEAEKAADESERGMKVLEN 132
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

```

QY      48 KEVQSIIVDDGAVDCEBRTGTSNYWAFPSKJLHARKHLKEVLSEQLSESSGOK-----HASTL 103
Db      133 R-----ALXDBEERMELOEI-----OLKEAHHAEEADRYEEVARKLTVVEGEI 176

QY      104 OKSIEKAKI--GRCTEERTRLAKELSLRDOEQJLKAVEREKYXCDPQVVEIRQANKY 161
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      177 ENTERRALNBSKC-----SLSEEELKYVTYNMKSLEQAEBKYSKEDRYEEIKVLTDK 231

QY      162 AKEAANRWTDNIFAIKSWAKRKGFEENKIDRTFGIPED 200
           |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      232 LKEAETRAE---FAERSVA-----KLKXTIDLLED 258

RESULT 12
US-09-914-259-49
; Sequence 49, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 284
; TYPE: prf
; ORGANISM: Rana temporaria
; US-09-914-259-49 .
```

```

Query Match          10.5%; Score 110; DB 4; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.0092;
Matches      48; Conservative   39; Mismatches    80; Indels   40; Gaps     8

QY      1 MSKKKGISAE--EKRTMMEIFSETKDFQ-----LKDLEKIAPK-EKGIYAMSV 47
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       73 LAEKKATPAADVASLNRRIGQVEBELDRAQRRLATALQCKLEAEAKADBSRGKVIEN 132
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      48 KEVLDSLVDDGMVDCERIGTSNYWAFPSKALHARKGLFVLESQSSESGQHASIQSI 107
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

DB      133 R----ALNDEKEIEIQEI-----QLXEAKHIAEBADRKYEEFAARLLVIIEGDL 176
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      108 EKAKIGRETER--TRLAKESLSLRDQROUKAEEVKYKDDPQVEVERIRQANKAYKA 165
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

DB      177 ERAE-ERAELSESCKALEBEELIKVTNNIKSLSEAQEKYSQKEDKYEETIKVLTDLKEA 235
           ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      166 ANRWTDNIFAIKSWARRKFGEFNKID 192
           |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

DB      236 ETR-----AEFARTVAKLKSID 254
           |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 13
US-09-438-185A-708
Sequence 708, Application US/09438185A
Patent No. 6822071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT FILING DATE: US/09/438,185A
PRIOR APPLICATION NUMBER: 2002-03-13
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1999-04-08

```

```

: NUMBER OF SEQ ID NOS: 1074
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 708
: LENGTH: 170
: TYPE: PRT
: ORGANISM: Chlamydia pneumoniae
: FEATURE:
: OTHER INFORMATION: CPN0706
: US-09-438-185A-708

Query Match          10.5%  Score 109.5; DB 4; Length 170;
Best Local Similarity 29.8%; Pred. No. 0.0052;
Matches 48; Conservative 29; Mismatches 51; Indels 33; Gaps 8

```

```

QY 10 EEKRTTMEIIFETPDVQLKDLKLTAKKEGKITAMSKVELOSLVDGNCVDERIGTSN 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 27 KEKR -RLTEIQE-----KLKREKARDKVNHHMQKIQQ -LRLLDEGTT----- 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 YWAPPSKTLHAKKHKLVEYLSQLS-----GSGKASL--QKSIKAKG--RCETEE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 -----SDAVLQIKYIKVAVQVLSSEBEKKNKQEVYLAASKLELGEAVNLKRRTEE 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 RTRFLAKELSLRDQRBQLKAEVEYKYCDPQVVEIRQANK 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 KTRLHKEWMKALKKEEARAE -EKQDMGGLRLPOLQOKKK 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 14
US-09-914-259--51
: Sequence 51, Application US/09914259
: Patent No. 649536
: GENERAL INFORMATION:
: APPLICANT: Makowski, Lee
: APPLICANT: Hyman, Paul
: APPLICANT: Williams, Mark
: TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
: FILE REFERENCE: 8471-010-999
: CURRENT APPLICATION NUMBER: US/09/914,259
: CURRENT FILING DATE: 2000-11-21
: NUMBER OF SEQ ID NOS: 180
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 51
: LENGTH: 264
: TYPE: PRT
: ORGANISM: Xenopus laevis
US-09-914-259-51

```

Query Match 10.4%; Score 109; DB 4; Length 284;
Best Local Similarity 23.7%; Pred. No. 0.011;
Matches 49; Conservative 37; Mismatches 81; Indels 40; Gaps 8

```
QY      1 MSKKKGLSAAE--EKRTMMEIFSEYKDFQ-----LKDLKTAAPK-EKGITAMSV 47
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      73 LSKDQATDAEDGVASLNRRIQLVVEEELDRAQERLSTALQKLEAEKKADESEBGMKVIEN 132

QY      48 KEVQSLVDGDMVDCERIGTSNYWPAFPSKRLHARKHLEVLSEQLSRSGQKHAQLQSKI 107
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      133 R---ALKDESKMELQET-----OLKEAKHIAESADRKYEVEAVKLVIIESDL 176

QY      108 EKAKIGRCETEER--TRLAKELSSLRQREQLKAVEKYKDCDPQVVEIRQANVAKEA 165
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      177 ERAAEKRAELSESKCALEELBLKTVTNLNLKSLAEQAERYSQKEDPYEIEIKVLTDKLEA 235

QY      166 ANRWTDNIFAIKSWAKRKFGPEEKID 192
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      236 ETR-----AEFAERTVAKLEKSID 254
```

RESULT 15
US-09-248-796A-20276
; Sequence 20276, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20276
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20276

Query Match 10.3%; Score 108; DB 4; Length 817;
Best Local Similarity 24.4%; Pred. No. 0.059;
Matches 49; Conservative 43; Mismatches 85; Indels 24; Gaps 8;
QY 9 ABEKTRMMEIFSETQDVDFQKLEKIAPEKIGITAMSVK--EVLQSLVD--DGMVDC 63
DB 426 ABEKQV-LDEKRENQDRIDTEAEKIAARKQELBEIQAKDEILKPTIDELKESAKLE 484
QY 64 RIGTSNYWYAFPSKALH---ARKHLEVLSQLSEGSQKHAISLQKSIKAKIGRCETEE 119
DB 485 EYTNARDELANEVYKASEDLNKEYEELKLELESKLQEAKNIDIEKYTTDIEEA-----TAK 538
QY 120 RPLALKELSLRQREQQLKAEVER-YKDCDPQVVEEIRQANKYAKAANRWTDNIFAIS 178
DB 539 HESTDEVAALQELHDEKADAEKEHEDLDGKLELEKQQLHLHEDKATKKDILAIIDE 598
QY 179 WAKRKFGEENKIDRTFGIPE 199
DB 599 KVK-----DEKINSE-LPE 612

Search completed: July 27, 2005, 09:42:31
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2005, 09:41:05 ; Search time 156 Seconds
(without alignments)
511.176 Million cell updates/sec

Title: US-10-087-190-3
Perfect score: 1047
Sequence: 1 MSKKKGLSAEKRRTMWEIF.....FENKIDRTPGIDPDYID 205

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1741741 seqs, 386992284 residues

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	205	US-09-799-250-2	Sequence 2, Appli.
2	1047	100.0	205	US-10-087-190-3	Sequence 3, Appli.
3	1047	100.0	205	US-10-087-190-14	Sequence 14, Appli
4	1047	100.0	205	US-10-087-190-20	Sequence 20, Appli
5	1047	100.0	205	US-10-087-190-21	Sequence 21, Appli
6	1047	100.0	205	US-10-087-190-22	Sequence 22, Appli
7	1047	100.0	205	US-10-087-190-61	Sequence 61, Appli
8	1047	100.0	205	US-10-408-765A-1821	Sequence 1821, Ap
9	1036.5	99.0	206	US-10-087-190-44	Sequence 44, Appli
10	1036.5	99.0	206	US-10-087-190-45	Sequence 45, Appli
11	1011	96.6	198	US-10-087-190-24	Sequence 24, Appli

12	975	93.1	190	US-10-087-190-13	Sequence 13, Appli
13	975	93.1	190	US-10-087-190-19	Sequence 19, Appli
14	975	93.1	190	US-10-087-190-67	Sequence 67, Appli
15	975	93.1	190	US-10-087-190-68	Sequence 68, Appli
16	975	93.1	190	US-10-087-190-69	Sequence 69, Appli
17	948	90.5	205	US-09-799-250-4	Sequence 4, Appli
18	948	90.5	205	US-10-087-190-23	Sequence 23, Appli
19	945.5	90.3	190	US-10-087-190-11	Sequence 11, Appli
20	945.5	90.3	190	US-10-087-190-18	Sequence 18, Appli
21	945.5	90.3	190	US-10-087-190-62	Sequence 62, Appli
22	945.5	90.3	190	US-10-087-190-63	Sequence 63, Appli
23	614	58.6	122	US-10-087-190-53	Sequence 53, Appli
24	591	56.4	122	US-10-087-190-9	Sequence 9, Appli
25	591	56.4	122	US-10-087-190-17	Sequence 17, Appli
26	591	56.4	122	US-10-087-190-54	Sequence 54, Appli
27	591	56.4	122	US-10-087-190-55	Sequence 55, Appli
28	584	55.8	119	US-10-087-190-7	Sequence 7, Appli
29	584	55.8	119	US-10-087-190-16	Sequence 16, Appli
30	584	55.8	119	US-10-087-190-47	Sequence 47, Appli
31	465	44.4	126	US-10-087-190-5	Sequence 5, Appli
32	465	44.4	126	US-10-087-190-15	Sequence 15, Appli
33	465	44.4	126	US-10-087-190-46	Sequence 46, Appli
34	434	41.5	207	US-10-425-115-253013	Sequence 253013,
35	427	40.8	207	US-10-437-963-162899	Sequence 162899,
36	394	37.6	235	US-10-437-963-162901	Sequence 162901,
37	349	33.3	200	US-09-799-250-5	Sequence 5, Appli
38	349	33.3	200	US-10-087-190-25	Sequence 25, Appli
39	287	27.4	79	US-10-106-698-5443	Sequence 5443, Appli
40	255.5	24.4	14	US-10-083-357-1304	Sequence 1304, Ap
41	214.5	20.5	74	US-10-424-599-192650	Sequence 192650,
42	121.5	11.6	715	US-10-408-765A-2096	Sequence 2096, Ap
43	121.5	11.6	1881	US-10-032-585-7646	Sequence 7646, Ap
44	120.5	11.5	1798	US-09-981-151A-48	Sequence 48, Appli
45	120	11.5	882	US-10-732-923-3300	Sequence 3300, Ap

ALIGNMENTS

```
RESULT 1
US-09-799-250-2
; Sequence 2, Application US/09799250
; Publication No. US20030032087A1
GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09799,250
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-2

Query Match      100.0%; Score 1047; DB 10; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.8e-78;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSKKKGLSAEKRRTMWEIFSETKDVFLKLEKAPAREKGTANSYKVLQSLVDENV 60
        |||
DB       1 MSKKKGLSAEKRRTMWEIFSETKDVFLKLEKAPAREKGTANSYKVLQSLVDENV 60
        |||
QY      61 DCERIGTSNYWAFPSKALHARKHKLVLLESQLSGSGKHASLQSIKAKIGRCETBEER 120
```

Db	61	DCERIGTSNYVAFPSKALHARKHKLVEISQSEBSQGHASIQXSIERAKIGRCETBER	120
Qy	121	TRLAELSSLRDQREOLKAVEREKKDCDPQVVEEIQANKVAKAANRWTDNIFAIKSWA	180
Db	121	TRLAELSSLRDQREOLKAVEREKKDCDPQVVEEIQANKVAKAANRWTDNIFAIKSWA	180
Qy	181	KRKGFEENKIDRTFGIPEDPYID	205
Db	181	KRKGFEENKIDRTFGIPEDPDYID	205

```

RESULT 2
US-10-087-190-3
; Sequence 3, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challisa-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raltano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-3

```

RESULT 3
 US-10-087-190-14
 ; Sequence 14, Application US/10087190
 ; Publication No. US20030223957A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Challite-Eld, Pia M.
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Paris, Mary
 ; APPLICANT: Afari, Daniel E. H.
 ; APPLICANT: Ge, Wangmao

```

; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087.190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-14

Query Match: 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.8e-78;
Match 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKRRTMEIIFSETKDVFLKLEIAIPKEKGTITANSVXEVLQSLVDDGV 60
DB 1 MSKKKGLSAEKRRTMEIIFSETKDVFLKLEIAIPKEKGTITANSVXEVLQSLVDDGV 60
QY 61 DCERGTSTNYWAPFSKALHARKHKLTVLESOLSGSHASLOKSIERAKIGRCETEE 120
DB 61 DCERGTSTNYWAPFSKALHARKHKLTVLESOLSGSHASLOKSIERAKIGRCETEE 120
QY 121 TRIAELISLRDREQLAEVEKYKDCPPOVEEIRQANKVAKEAANRTDNI PAIKSWA 180
DB 121 TRIAELISLRDREQLAEVEKYKDCPPOVEEIRQANKVAKEAANRTDNI PAIKSWA 180
QY 181 KRKGFEEENKIDRTFGIPEDDFYID 205
DB 181 KRKGFEEENKIDRTFGIPEDDFYID 205

```

```

RESULT 4
US-10-087-190-20
; Sequence 20, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Ageneys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 205
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-087-190-20

```

```

Query Match 100.0% Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3,88-78;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 MSKKKGSLAERKTRPMELFSETKDVFOLOKLEKAPAEKGI7TMSVKEVTLQSLVDDGV 60
DB      1 MSKKKGSLAERKTRPMELFSETKDVFOLOKLEKAPAEKGI7TMSVKEVTLQSLVDDGV 60

```

Qy 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIKAKIGRCETEE 120
Db 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIKAKIGRCETEE 120
Qy 121 TRLAELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKAANRWTNIFAIKSWA 180
Db 121 TRLAELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKAANRWTNIFAIKSWA 180
Qy 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 5
US-10-087-190-21
Sequence 21, Application US/10087190
Publication No. US20030223997A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Challita-Bid, Pia M.
APPLICANT: Hubert, Rene S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Afar, Daniel E. H.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OF INVENTION: ENTITLED 121P11 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 205
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-21

Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.8e-78;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGLSAEKRTTMEIFSETKOVFQDKLEKIAPEKGTITMSVKEVLSLVDDGMV 60
Db 1 MSKKKGLSAEKRTTMEIFSETKOVFQDKLEKIAPEKGTITMSVKEVLSLVDDGMV 60
Qy 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIKAKIGRCETEE 120
Db 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIKAKIGRCETEE 120
Qy 121 TRLAELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKAANRWTNIFAIKSWA 180
Db 121 TRLAELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKAANRWTNIFAIKSWA 180
Qy 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 6
US-10-087-190-22
Sequence 22, Application US/10087190
Publication No. US20030223997A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Challita-Bid, Pia M.
APPLICANT: Hubert, Rene S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Afar, Daniel E. H.

APPLICANT: Ge, Wangmao
APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OF INVENTION: ENTITLED 121P11 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 205
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-22

Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.8e-78;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGLSAEKRTTMEIFSETKOVFQDKLEKIAPEKGTITMSVKEVLSLVDDGMV 60
Db 1 MSKKKGLSAEKRTTMEIFSETKOVFQDKLEKIAPEKGTITMSVKEVLSLVDDGMV 60
Qy 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIKAKIGRCETEE 120
Db 61 DCRIGTSNYWAFPSKALHARKHLEVLSEQLSEGSQGHASLOKSIKAKIGRCETEE 120
Qy 121 TRLAELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKAANRWTNIFAIKSWA 180
Db 121 TRLAELSLRQREQLKAEVEKYDCDPQVVEIRQANKVAKAANRWTNIFAIKSWA 180
Qy 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 7
US-10-087-190-61
Sequence 61, Application US/10087190
Publication No. US20030223997A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Challita-Bid, Pia M.
APPLICANT: Hubert, Rene S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Afar, Daniel E. H.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OF INVENTION: ENTITLED 121P11 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 61
LENGTH: 205
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-61

Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.8e-78;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGLSAEKRTTMEIFSETKOVFQDKLEKIAPEKGTITMSVKEVLSLVDDGMV 60
Db 1 MSKKKGLSAEKRTTMEIFSETKOVFQDKLEKIAPEKGTITMSVKEVLSLVDDGMV 60

Db 61 DCEIGTSNYWAFPSKALHARKHLEVESQESQKHSLOKSEKAKIGRCETEE 120
QY 120 RRLAKESLSDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSW 179
Db 121 RRLAKESLSDQREQLKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSW 180
QY 180 AKRKGFEENKIDRTFGIPEDFDYID 205
Db 181 AKRKGFEENKIDRTFGIPEDFDYID 206

RESULT 11

US-10-087-190-24
Sequence 24, Application US/10087190
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Challita-Bid, Pia M.
APPLICANT: Hubert, Rene S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Afar, Daniel E. H.
APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 198
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-24

Query Match 96.6%; Score 1011; DB 15; Length 198;
Best Local Similarity 100.0%; Pred. No. 3.3e-75;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KGLSABEKRTRMELFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCER 64
Db 1 KGLSABEKRTRMELFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCER 60
QY 65 IGTSNYWAFPSKALHARKHLEVESQESQKHSLOKSEKAKIGRCETEE 124
Db 61 IGTSNYWAFPSKALHARKHLEVESQESQKHSLOKSEKAKIGRCETEE 120
QY 125 KGLSABEKRTRMELFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCER 184
Db 121 KGLSABEKRTRMELFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCER 180
QY 185 GFEENKIDRTFGIPEDFD 202
Db 181 GFEENKIDRTFGIPEDFD 198

RESULT 12

US-10-087-190-13
Sequence 13, Application US/10087190
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Challita-Bid, Pia M.
APPLICANT: Hubert, Rene S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Afar, Daniel E. H.
APPLICANT: Ge, Wangmao

APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 190
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-13

Query Match 93.1%; Score 975; DB 15; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.9e-72;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEIFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCERIGTSNYWAF 75
Db 1 MMEIFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCERIGTSNYWAF 60
QY 76 SKALHARKHLEVESQESQKHSLOKSEKAKIGRCETEE 135
Db 61 SKALHARKHLEVESQESQKHSLOKSEKAKIGRCETEE 120
QY 136 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 195
Db 121 QKAEVEKYKDCDPQVEEIRQANKVAKAANRWTDNIFAIKSWAKRKGFEENKIDRTF 180
QY 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

RESULT 13

US-10-087-190-19
Sequence 19, Application US/10087190
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Challita-Bid, Pia M.
APPLICANT: Hubert, Rene S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Farris, Mary
APPLICANT: Afar, Daniel E. H.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
CURRENT FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 190
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-19

Query Match 93.1%; Score 975; DB 15; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.9e-72;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEIFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCERIGTSNYWAF 75
Db 1 MMEIFSEKDVFOLEKIPKKEGITAMSVKYLQSLVDDGAVDCERIGTSNYWAF 60

```

QY      76  SKALHARKHKEVLESQLSSEGSOKJASLOKSIEKAKIGCEFEERTPLAKELSSLRDRE 135
      |||||||
Db      61  SKALHARKHKEVLESQLSSEGSOKJASLOKSIEKAKIGCEFEERTPLAKELSSLRDRE 120
      |||||||

QY      136  QLKAEVEKYKDCDPQVVEEIRQANKVKAENRMTDNIFAIKSMARKKPFEEENIDRTF 195
      |||||||
Db      121  QLKAEVEKYKDCDPQVVEEIRQANKVKAENRMTDNIFAIKSMARKKPFEEENIDRTF 180
      |||||||

QY      196  GIPEDFDYID 205
      |||||||
Db      181  GIPEDFDYID 190
      |||||||

RESULT 14
US-10-087-190-67
; Sequence 67, Application US/10087190
; Publication No. US2003022397A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-087-190-67

```

RESULT 15
 US-10-087-190-68
 : Sequence 68, Application US/100871900
 : Publication No. US2003023997A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Agensys, Inc.
 : APPLICANT: Chailita-Eld, Pia M.
 : APPLICANT: Hubbert, Rene S.
 : APPLICANT: Raitano, Arthur B.
 : APPLICANT: Paris, Mary
 : APPLICANT: Afar, Daniel E. H.

```

APPLICANT: Ge, Wangmao
APPLICANT: Jakobovitz, Aya
TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
TITLE OR INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
FILE REFERENCE: 51158-20034.20
CURRENT APPLICATION NUMBER: US/10/087,190
PRIOR FILING DATE: 2003-01-28
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 68
LENGTH: 190
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-087-190-68

93.1%; Score 975; DB 15; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.9e-72;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEFSTKCVFQKDLKELAPKRGKITAMVKEVLSQNDGMDRCRIGSNYYMAP 75
Db 1 MMEFSTKCVFQKDLKELAPKRGKITAMVKEVLSQNDGMDRCRIGSNYYMAP 60
QY 76 SKALHARKHLELVESQSEGSQGHASLOKSIIEKAKICRCETERTLAKELSLRDQRE 135
Db 61 SKALHARKHLELVESQSEGSQGHASLOKSIIEKAKICRCETERTLAKELSLRDQRE 120
QY 136 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAIKSMARKRGPEENKIDRTF 195
Db 121 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAIKSMARKRGPEENKIDRTF 180
QY 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

```

Search completed: July 27, 2005, 09:54:35
Job time : 157 secs